heart
brain
placenta
lung
liver
skeletal muscle
kidney
pancreas
stomach
thyroid
spinal cord
lymph node
trachea
adrenal gland
bone marrow
spleen
thymus
prostate
testis
ovary
small intestine
colon
peripheral blood leukocyte

Fig. 1. Northern analysis of hepsin mRNA in human tissues. Northern hybridization was performed with multiple human tissues blots (CLONTECH) using a full-length human hepsin cDNA. Hepsin mRNA was detected in samples from liver, kidney, pancreas, stomach, thyroid, prostate and testis.

**Relative Level of Hepsin
mRNA Expression**
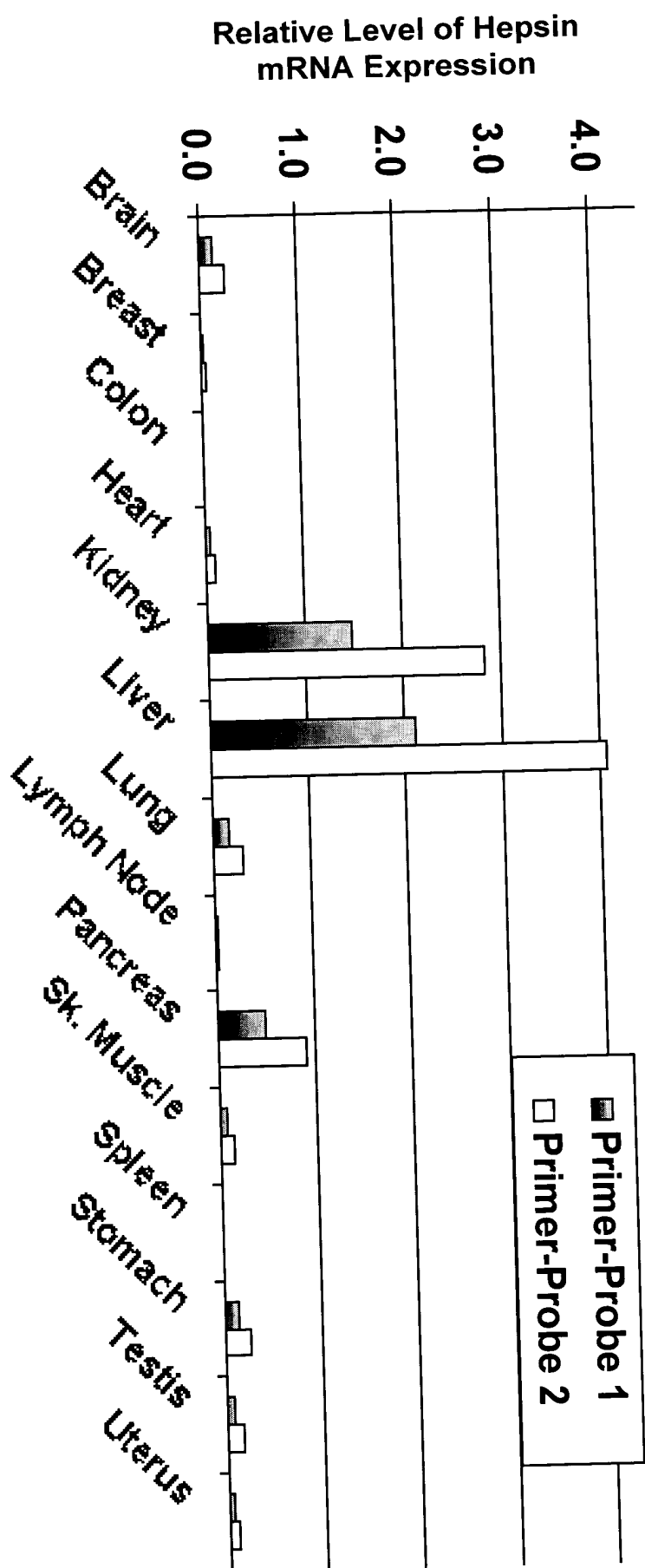
Fig. 2. Taqman analysis of hepsin mRNA expression in human tissues. Taqman-based quantitative PCR analysis of hepsin mRNA expression was performed using two sets of hepsin-specific primers (Primer-Probes 1 and 2). High levels of hepsin mRNA expression were detected in samples from liver and kidney.
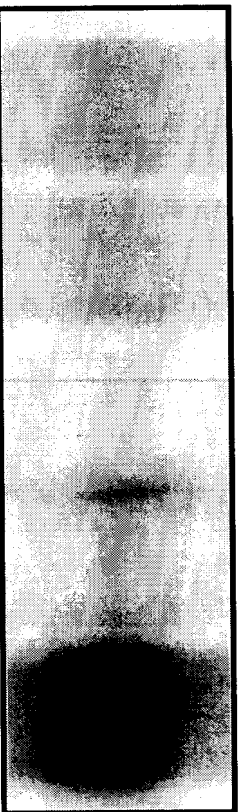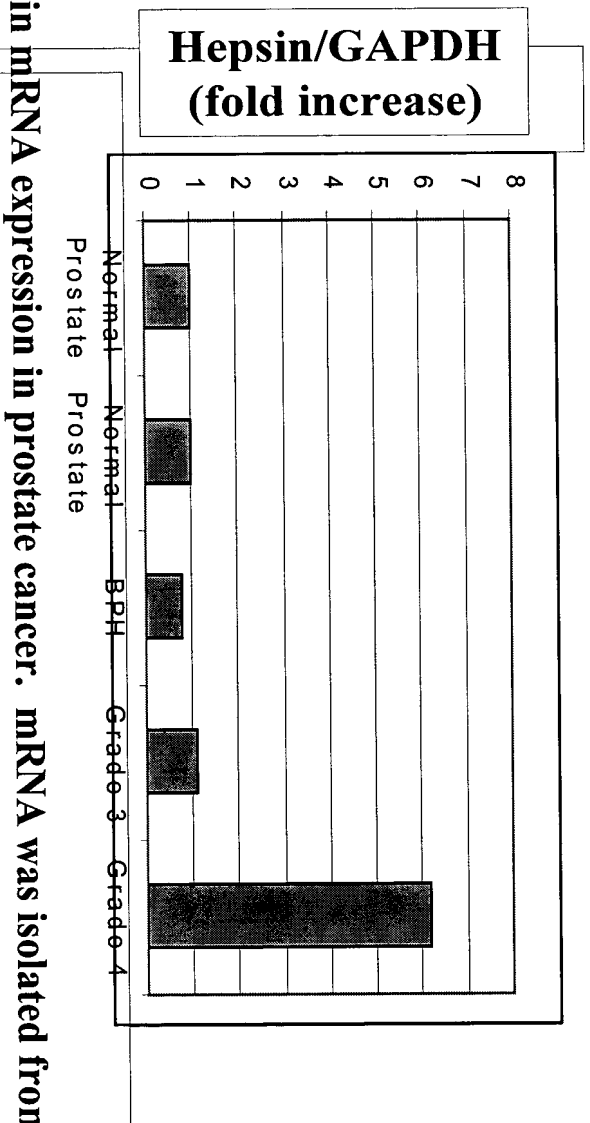
Fig. 3. Hepsin mRNA expression in prostate cancer. mRNA was isolated from normal prostate, BPH and prostate cancer tissues. Northern blot was prepared and hybridized with a human hepsin cDNA probe. Hepsin mRNA expressed was significantly higher in an advanced prostate cancer sample. The ratio of hepsin/GAPDH mRNA was ~6-fold higher in the advanced cancer than in normal prostate, BPH or primary prostate cancer.

Relative Expression of Hepsin mRNA

Legend:
- ■ Primer-Probe 1
- ▨ Primer-Probe 2

X-axis: 0.00, 0.50, 1.00, 1.50, 2.00, 2.50, 3.00

Categories: LNCaP, BPH, BPH, Early Stage, Early Stage, Early Stage, Early Stage, Late Stage, Late Stage, Late Stage
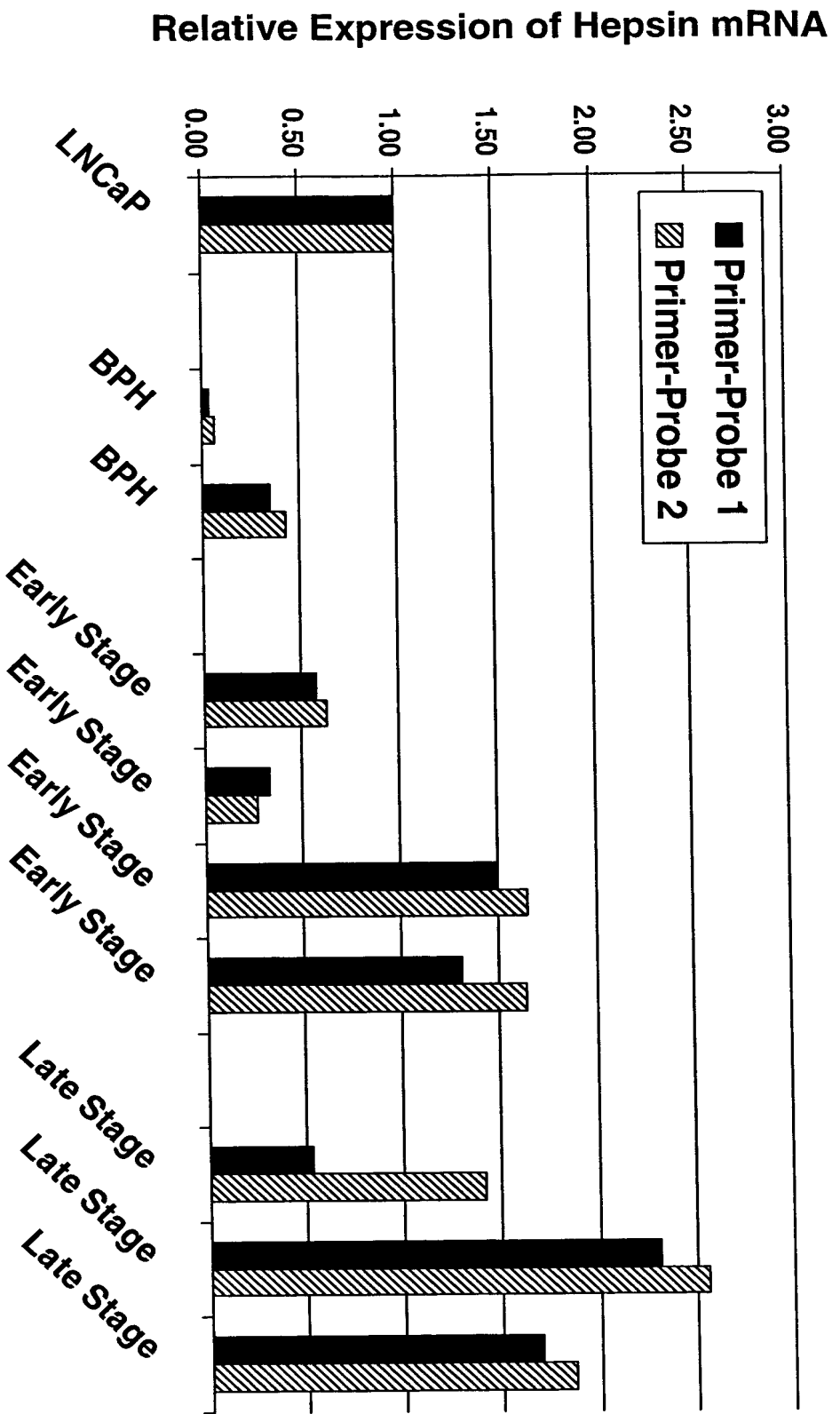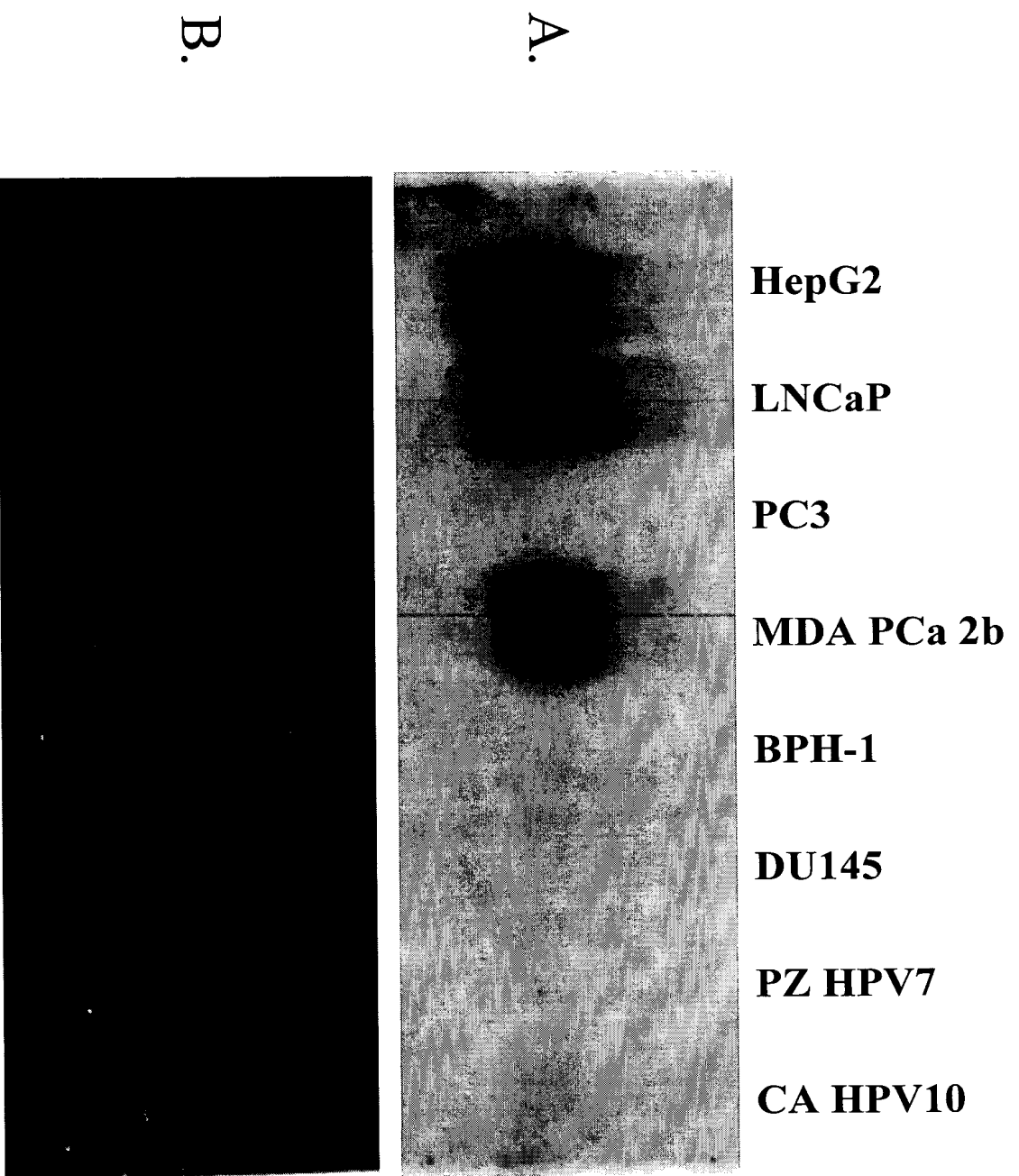
Figure 4. Taqman analysis of hepsin mRNA expression in prostate tissue. mRNA samples were prepared from LNCaP cells, BPH and prostate cancer (Gleason grade 3 and 4) tissues. Quantitative PCR analysis was performed using two set of primers specific for human hepsin (Primer-Probes 1 and 2). Modest to more dramatic increases of hepsin mRNA expression were detected in prostate cancer samples.

A.

HepG2

LNCaP

PC3

MDA PCa 2b

BPH-1

DU145

PZ HPV7

CA HPV10

B.

Fig. 5. Northern analysis of hepsin mRNA expression in prostate-derived cell lines. mRNA was isolated from prostate-derived cell lines, LNCaP, PC3, MDA PCa2b, BPH-1, DU145, PZ HPV7, CA HPV10 and a control cell line, HepG2. Northern analysis detected hepsin mRNA expression in LNCaP, MDA Pca 2b and HepG2 cells (upper panel). As a control, the agarose gel was stained with ethdium bromide to show equal sample loading.

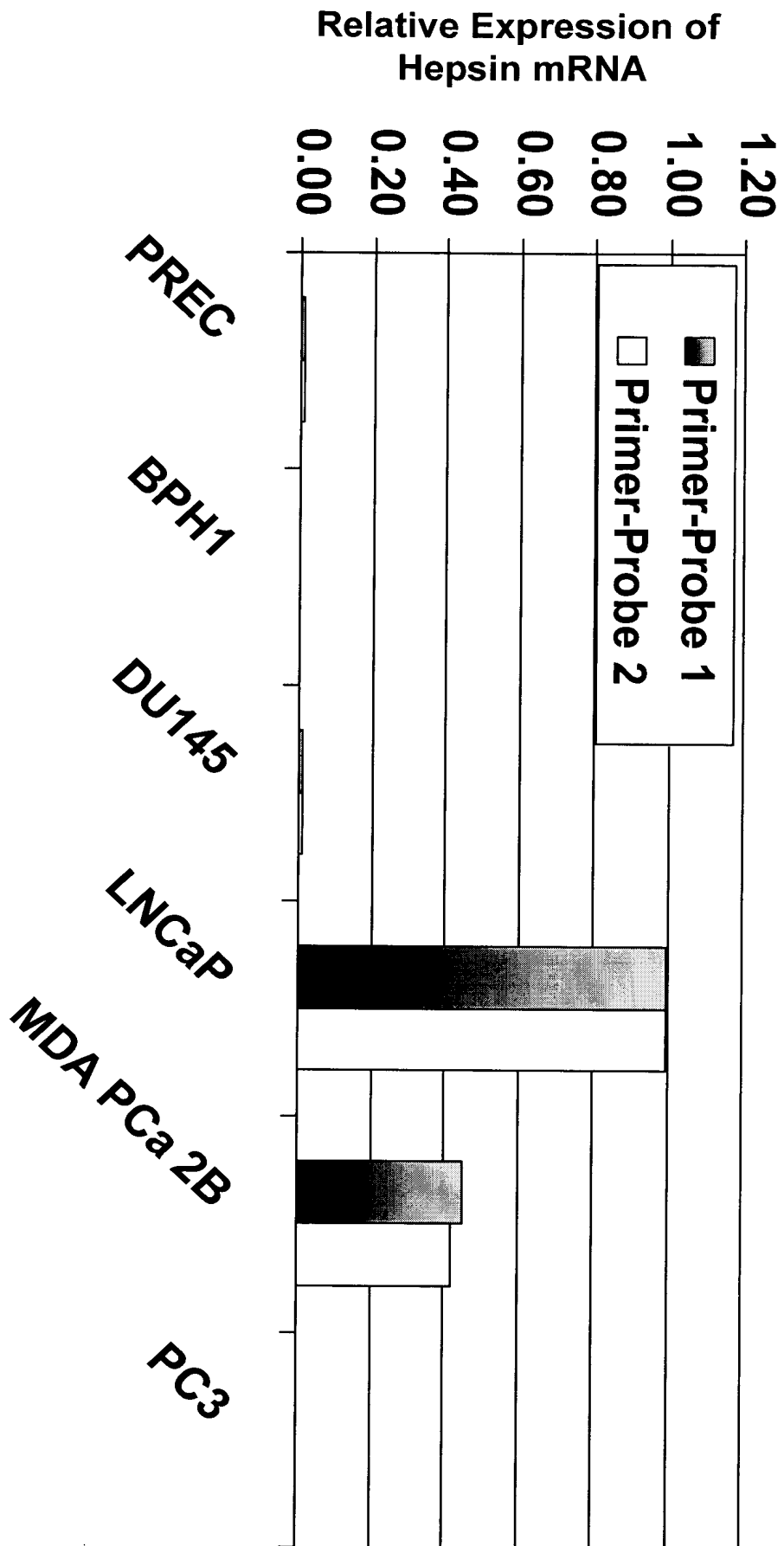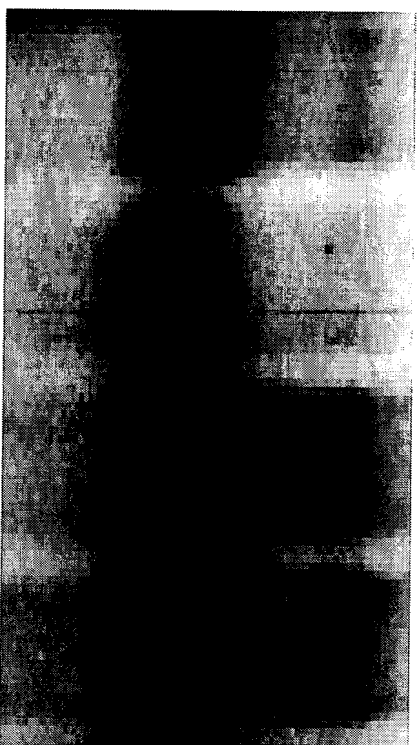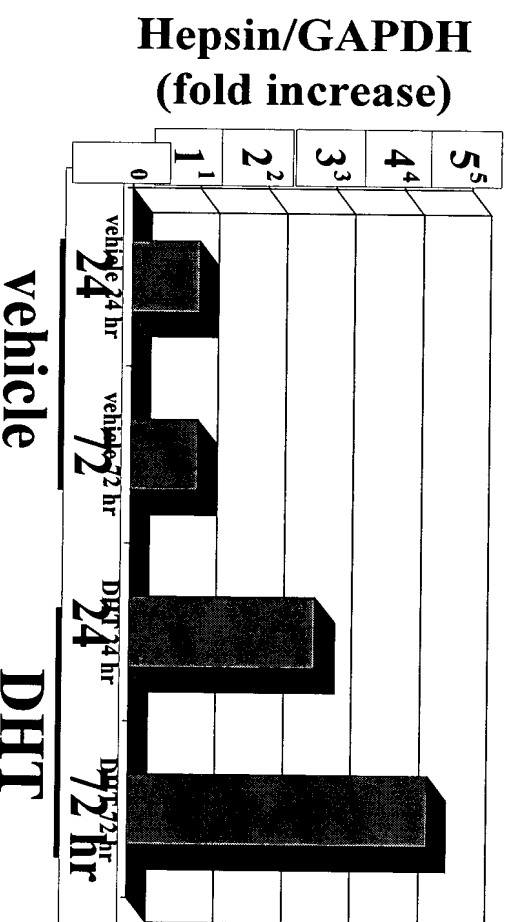**Relative Expression of Hepsin mRNA**

Fig. 6. Taqman analysis of hepsin mRNA expression in prostate-derived cell lines. mRNA was isolated from prostate-derived cell lines. Taqman analysis was performed using two sets of hepsin specific primers (Primer-Probes 1 and 2). Hepsin mRNA expression was detected in LNCaP and MDA Pca 2B cells.

A.

Hepsin

B.

Hepsin/GAPDH (fold increase)

vehicle   DHT

vehicle 24 hr | vehicle 72 hr | DHT 24 hr | DHT 72 hr

**Fig. 7.** Up-regulation of hepsin mRNA expression in LNCaP cells by dihydrotestosterone (DHT). LNCaP cells were treated with DHT (10 nM) and harvested at indicted time points. Total RNA samples were isolated. Northern blots were hybridized with a human hepsin cDNA probe and re-probed with a GAPDH cDNA. Hepsin mRNA expression was significantly enhanced in LNCaP cells treated with DHT for 24 and 72 hours.
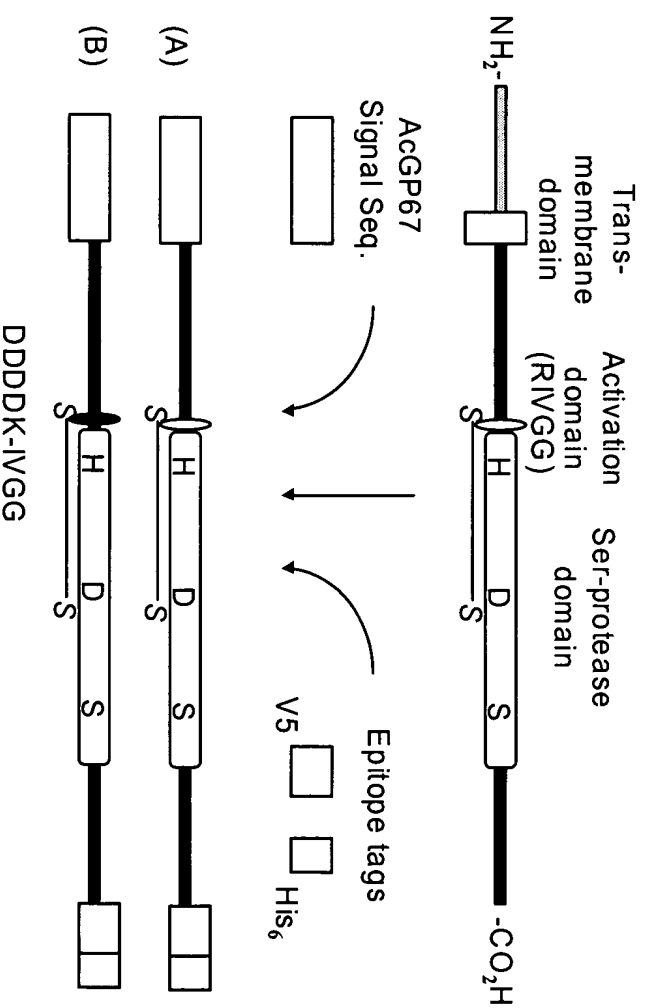
Fig. 8. Expression of soluble wild-type and mutant hepsins. A schematic presentation of human hepsin protein is shown on top. To express soluble wild-type hepsin (A), the extracellular domain of hepsin is cloned into vector pAcGP67 that contains a sequence coding for a signal peptide at the 5' end and a sequence coding for a V5 epitope and a His6 tag at the 3' end. In construct B, a sequence coding for an enterokinase recognition peptide, DDDDK, is inserted to replace the coding sequence for Arg 162. The mutation will allow hepsin to be activated by enterokinase.

FIG 9 (1-7) pIRESpuro2W/hepEK_k

```
   1   GACGGATCGG  GAGATCTCCC  GATCCCCTAT  GGTCGACTCT  CAGTACAATC
       CTGCCTAGCC  CTCTAGAGGG  CTAGGGGATA  CCAGCTGAGA  GTCATGTTAG
  51   TGCTCTGATG  CCGCATAGTT  AAGCCAGTAT  CTGCTCCCTG  CTTGTGTGTT
       ACGAGACTAC  GGCGTATCAA  TTCGGTCATA  GACGAGGGAC  GAACACACAA
 101   GGAGGTCGCT  GAGTAGTGCG  CGAGCAAAAT  TTAAGCTACA  ACAAGGCAAG
       CCTCCAGCGA  CTCATCACGC  GCTCGTTTTA  AATTCGATGT  TGTTCCGTTC
 151   GCTTGACCGA  CAATTGCATG  AAGAATCTGC  TTAGGGTTAG  GCGTTTTGCG
       CGAACTGGCT  GTTAACGTAC  TTCTTAGACG  AATCCCAATC  CGCAAAACGC
 201   CTGCTTCGCG  ATGTACGGGC  CAGATATACG  CGTTGACATT  GATTATTGAC
       GACGAAGCGC  TACATGCCCG  GTCTATATGC  GCAACTGTAA  CTAATAACTG
 251   TAGTTATTAA  TAGTAATCAA  TTACGGGGTC  ATTAGTTCAT  AGCCCATATA
       ATCAATAATT  ATCATTAGTT  AATGCCCCAG  TAATCAAGTA  TCGGGTATAT
 301   TGGAGTTCCG  CGTTACATAA  CTTACGGTAA  ATGGCCCGCC  TGGCTGACCG
       ACCTCAAGGC  GCAATGTATT  GAATGCCATT  TACCGGGCGG  ACCGACTGGC
 351   CCCAACGACC  CCCGCCCATT  GACGTCAATA  ATGACGTATG  TTCCCATAGT
       GGGTTGCTGG  GGGCGGGTAA  CTGCAGTTAT  TACTGCATAC  AAGGGTATCA
 401   AACGCCAATA  GGGACTTTCC  ATTGACGTCA  ATGGGTGGAC  TATTTACGGT
       TTGCGGTTAT  CCCTGAAAGG  TAACTGCAGT  TACCCACCTG  ATAAATGCCA
 451   AAACTGCCCA  CTTGGCAGTA  CATCAAGTGT  ATCATATGCC  AAGTACGCCC
       TTTGACGGGT  GAACCGTCAT  GTAGTTCACA  TAGTATACGG  TTCATGCGGG
 501   CCTATTGACG  TCAATGACGG  TAAATGGCCC  GCCTGGCATT  ATGCCCAGTA
       GGATAACTGC  AGTTACTGCC  ATTTACCGGG  CGGACCGTAA  TACGGGTCAT
 551   CATGACCTTA  TGGGACTTTC  CTACTTGGCA  GTACATCTAC  GTATTAGTCA
       GTACTGGAAT  ACCCTGAAAG  GATGAACCGT  CATGTAGATG  CATAATCAGT
 601   TCGCTATTAC  CATGGTGATG  CGGTTTTGGC  AGTACATCAA  TGGGCGTGGA
       AGCGATAATG  GTACCACTAC  GCCAAAACCG  TCATGTAGTT  ACCCGCACCT
 651   TAGCGGTTTG  ACTCACGGGG  ATTTCCAAGT  CTCCACCCCA  TTGACGTCAA
       ATCGCCAAAC  TGAGTGCCCC  TAAAGGTTCA  GAGGTGGGGT  AACTGCAGTT
 701   TGGGAGTTTG  TTTTGGCACC  AAAATCAACG  GGACTTTCCA  AAATGTCGTA
       ACCCTCAAAC  AAAACCGTGG  TTTTAGTTGC  CCTGAAAGGT  TTTACAGCAT
 751   ACAACTCCGC  CCCATTGACG  CAAATGGGCG  GTAGGCGTGT  ACGGTGGGAG
       TGTTGAGGCG  GGGTAACTGC  GTTTACCCGC  CATCCGCACA  TGCCACCCTC
 801   GTCTATATAA  GCAGAGCTCT  CTGGCTAACT  AGAGAACCCA  CTGCTTACTG
       CAGATATATT  CGTCTCGAGA  GACCGATTGA  TCTCTTGGGT  GACGAATGAC
 851   GCTTATCGAA  ATTAATACGA  CTCACTATAG  GGAGACCCAA  GCTTGGTACC
       CGAATAGCTT  TAATTATGCT  GAGTGATATC  CCTCTGGGTT  CGAACCATGG
  +3                                            M  E  T  D  T  L  L  W
 901   GAGCTCGGAT  CGATATCGCC  ACCATGGAGA  CAGACACACT  CCTGCTATGG
       CTCGAGCCTA  GCTATAGCGG  TGGTACCTCT  GTCTGTGTGA  GGACGATACC
  +3   V  L  L   L  W  V  P   G  S  T  G   D  A  P   D  R  S·
 951   GTACTGCTGC  TCTGGGTTCC  AGGTTCCACT  GGTGACGCTC  CGGACAGGAG
       CATGACGACG  AGACCCAAGG  TCCAAGGTGA  CCACTGCGAG  GCCTGTCCTC
  +3   ·S  D  Q   E  P  L  Y   P  V  Q  V   S  S  A   D  A  R  L
1001   TGACCAGGAG  CCGCTGTACC  CAGTGCAGGT  CAGCTCTGCG  GACGCTCGGC
       ACTGGTCCTC  GGCGACATGG  GTCACGTCCA  GTCGAGACGC  CTGCGAGCCG
```

FIGURE 9-1

```
+3    ·L  M   V   F   D  K   T   E   G   T   W   R   L   L   C   S   S
1051  TCATGGTCTT TGACAAGACG GAAGGGACGT GGCGGCTGCT GTGCTCCTCG
      AGTACCAGAA ACTGTTCTGC CTTCCCTGCA CCGCCGACGA CACGAGGAGC
+3     R   S   N   A   R   V   A   G   L   S   C   E   E   M   G   F  L·
1101  CGCTCTAACG CCAGGGTAGC CGGACTCAGC TGCGAGGAGA TGGGCTTCCT
      GCGAGATTGC GGTCCCATCG GCCTGAGTCG ACGCTCCTCT ACCCGAAGGA
+3    ·L  R   A   L   T   H   S   E   L   D   V   R   T   A   G   A  N  G
1151  CAGGGCACTG ACCCACTCCG AGCTGGACGT GCGAACGGCG GGCGCCAATG
      GTCCCGTGAC TGGGTGAGGC TCGACCTGCA CGCTTGCCGC CCGCGGTTAC
+3    ·G  T   S   G   F   F   C   V   D   E   G   R   L   P   H   T   Q
1201  GCACGTCGGG CTTCTTCTGT GTGGACGAGG GGAGGCTGCC CCACACCCAG
      CGTGCAGCCC GAAGAAGACA CACCTGCTCC CCTCCGACGG GGTGTGGGTC
+3     R   L   L   E   V   I   S   V   C   D   C   P   R   G   R   F  L·
1251  AGGCTGCTGG AGGTCATCTC CGTGTGTGAT TGCCCCAGAG GCCGTTTCTT
      TCCGACGACC TCCAGTAGAG GCACACACTA ACGGGGTCTC CGGCAAAGAA
+3    ·L  A   A   I   C   Q   D   C   G   R   R   K   L   P   V   D  D  D
1301  GGCCGCCATC TGCCAAGACT GTGGCCGCAG GAAGCTGCCC GTGGACGACG
      CCGGCGGTAG ACGGTTCTGA CACCGGCGTC CTTCGACGGG CACCTGCTGC
+3    ·D  D   K   I   V   G   G   R   D   T   S   L   G   R   W   P   W
1351  ACGACAAGAT CGTGGGAGGC CGGGACACCA GCTTGGGCCG GTGGCCGTGG
      TGCTGTTCTA GCACCCTCCG GCCCTGTGGT CGAACCCGGC CACCGGCACC
+3     Q   V   S   L   R   Y   D   G   A   H   L   C   G   G   S   L  L·
1401  CAAGTCAGCC TTCGCTATGA TGGAGCACAC CTCTGTGGGG GATCCCTGCT
      GTTCAGTCGG AAGCGATACT ACCTCGTGTG GAGACACCCC CTAGGGACGA
+3    ·L  S   G   D   W   V   L   T   A   A   H   C   F   P   E   R  N  R
1451  CTCCGGGGAC TGGGTGCTGA CAGCCGCCCA CTGCTTCCCG GAGCGGAACC
      GAGGCCCCTG ACCCACGACT GTCGGCGGGT GACGAAGGGC CTCGCCTTGG
+3    ·R  V   L   S   R   W   R   V   F   A   G   A   V   A   Q   A   S
1501  GGGTCCTGTC CCGATGGCGA GTGTTTGCCG GTGCCGTGGC CCAGGCCTCT
      CCCAGGACAG GGCTACCGCT CACAAACGGC CACGGCACCG GGTCCGGAGA
+3     P   H   G   L   Q   L   G   V   Q   A   V   V   Y   H   G   G  Y·
1551  CCCCACGGTC TGCAGCTGGG GGTGCAGGCT GTGGTCTACC ACGGGGGCTA
      GGGGTGCCAG ACGTCGACCC CCACGTCCGA CACCAGATGG TGCCCCCGAT
+3    ·Y  L   P   F   R   D   P   N   S   E   E   N   S   N   D   I  A  L
1601  TCTTCCCTTT CGGGACCCCA ACAGCGAGGA GAACAGCAAC GATATTGCCC
      AGAAGGGAAA GCCCTGGGGT TGTCGCTCCT CTTGTCGTTG CTATAACGGG
+3    ·L  V   H   L   S   S   P   L   P   L   T   E   Y   I   Q   P   V
1651  TGGTCCACCT CTCCAGTCCC CTGCCCCTCA CAGAATACAT CCAGCCTGTG
      ACCAGGTGGA GAGGTCAGGG GACGGGGAGT GTCTTATGTA GGTCGGACAC
+3     C   L   P   A   A   G   Q   A   L   V   D   G   K   I   C   T  V·
1701  TGCCTCCCAG CTGCCGGCCA GGCCCTGGTG GATGGCAAGA TCTGTACCGT
      ACGGAGGGTC GACGGCCGGT CCGGGACCAC CTACCGTTCT AGACATGGCA
+3    ·V  T   G   W   G   N   T   Q   Y   Y   G   Q   Q   A   G   V  L  Q
1751  GACGGGCTGG GGCAACACGC AGTACTATGG CCAACAGGCC GGGGTACTCC
      CTGCCCGACC CCGTTGTGCG TCATGATACC GGTTGTCCGG CCCCATGAGG
```

FIGURE 9-2

2

+3   ·Q   E   A   R   V   P   I   I   S   N   D   V   C   N   G   A   D

1801 AGGAGGCTCG AGTCCCCATA ATCAGCAATG ATGTCTGCAA TGGCGCTGAT
      TCCTCCGAGC TCAGGGGTAT TAGTCGTTAC TACAGACGTT ACCGCGACTA

+3   F   Y   G   N   Q   I   K   P   K   M   F   C   A   G   Y   P   E·

1851 TTCTATGGAA ACCAGATCAA GCCCAAGATG TTCTGTGCTG GCTACCCCGA
      AAGATACCTT TGGTCTAGTT CGGGTTCTAC AAGACACGAC CGATGGGGCT

+3   ·E   G   G   I   D   A   C   Q   G   D   S   G   G   P   F   V   C   E

1901 GGGTGGCATT GATGCCTGCC AGGGCGACAG CGGTGGTCCC TTTGTGTGTG
      CCCACCGTAA CTACGGACGG TCCCGCTGTC GCCACCAGGG AAACACACAC

+3   ·E   D   S   I   S   R   T   P   R   W   R   L   C   G   I   V   S

1951 AGGACAGCAT CTCTCGGACG CCACGTTGGC GGCTGTGTGG CATTGTGAGT
      TCCTGTCGTA GAGAGCCTGC GGTGCAACCG CCGACACACC GTAACACTCA

+3   W   G   T   G   C   A   L   A   Q   K   P   G   V   Y   T   K   V·

2001 TGGGGCACTG GCTGTGCCCT GGCCCAGAAG CCAGGCGTCT ACACCAAAGT
      ACCCCGTGAC CGACACGGGA CCGGGTCTTC GGTCCGCAGA TGTGGTTTCA

+3   ·V   S   D   F   R   E   W   I   F   Q   A   I   K   T   H   S   E   A·

2051 CAGTGACTTC CGGGAGTGGA TCTTCCAGGC CATAAAGACT CACTCCGAAG
      GTCACTGAAG GCCCTCACCT AGAAGGTCCG GTATTTCTGA GTGAGGCTTC

+3   ·A   S   G   M   V   T   Q   L   E   F   G   K   P   I   P   N   P

2101 CCAGCGGCAT GGTGACCCAG CTCGAATTCG GTAAGCCTAT CCCTAACCCT
      GGTCGCCGTA CCACTGGGTC GAGCTTAAGC CATTCGGATA GGGATTGGGA

+3   L   L   G   L   D   S   T   R   T   G   H   H   H   H   H   H

2151 CTCCTCGGTC TCGATTCTAC GCGTACCGGT CATCATCACC ATCACCATTG
      GAGGAGCCAG AGCTAAGATG CGCATGGCCA GTAGTAGTGG TAGTGGTAAC

2201 AGTTTAAAGC GGCCGCATAG ATAACTGATC CAGTGTGCTG GAATTAATTC
      TCAAATTTCG CCGGCGTATC TATTGACTAG GTCACACGAC CTTAATTAAG

2251 GCTGTCTGCG AGGGCCAGCT GTTGGGGTGA GTACTCCCTC TCAAAAGCGG
      CGACAGACGC TCCCGGTCGA CAACCCCACT CATGAGGGAG AGTTTTCGCC

2301 GCATGACTTC TGCGCTAAGA TTGTCAGTTT CCAAAAACGA GGAGGATTTG
      CGTACTGAAG ACGCGATTCT AACAGTCAAA GGTTTTTGCT CCTCCTAAAC

2351 ATATTCACCT GGCCCGCGGT GATGCCTTTG AGGGTGGCCG CGTCCATCTG
      TATAAGTGGA CCGGGCGCCA CTACGGAAAC TCCCACCGGC GCAGGTAGAC

2401 GTCAGAAAAG ACAATCTTTT TGTTGTCAAG CTTGAGGTGT GGCAGGCTTG
      CAGTCTTTTC TGTTAGAAAA ACAACAGTTC GAACTCCACA CCGTCCGAAC

2451 AGATCTGGCC ATACACTTGA GTGACAATGA CATCCACTTT GCCTTTCTCT
      TCTAGACCGG TATGTGAACT CACTGTTACT GTAGGTGAAA CGGAAAGAGA

2501 CCACAGGTGT CCACTCCCAG GTCCAACTGC AGGTCGAGCA TGCATCTAGG
      GGTGTCCACA GGTGAGGGTC CAGGTTGACG TCCAGCTCGT ACGTAGATCC

2551 GCGGCCAATT CCGCCCCTCT CCCTCCCCCC CCCCTAACGT TACTGGCCGA
      CGCCGGTTAA GGCGGGGAGA GGGAGGGGGG GGGGATTGCA ATGACCGGCT

2601 AGCCGCTTGG AATAAGGCCG GTGTGCGTTT GTCTATATGT GATTTTCCAC
      TCGGCGAACC TTATTCCGGC CACACGCAAA CAGATATACA CTAAAAGGTG

2651 CATATTGCCG TCTTTTGGCA ATGTGAGGGC CCGGAAACCT GGCCCTGTCT
      GTATAACGGC AGAAAACCGT TACACTCCCG GGCCTTTGGA CCGGGACAGA

2701 TCTTGACGAG CATTCCTAGG GGTCTTTCCC CTCTCGCCAA AGGAATGCAA
      AGAACTGCTC GTAAGGATCC CCAGAAAGGG GAGAGCGGTT TCCTTACGTT

FIGURE 9-3

3

```
2751  GGTCTGTTGA ATGTCGTGAA GGAAGCAGTT CCTCTGGAAG CTTCTTGAAG
      CCAGACAACT TACAGCACTT CCTTCGTCAA GGAGACCTTC GAAGAACTTC
2801  ACAAACAACG TCTGTAGCGA CCCTTTGCAG GCAGCGGAAC CCCCCACCTG
      TGTTTGTTGC AGACATCGCT GGGAAACGTC CGTCGCCTTG GGGGGTGGAC
2851  GCGACAGGTG CCTCTGCGGC CAAAAGCCAC GTGTATAAGA TACACCTGCA
      CGCTGTCCAC GGAGACGCCG GTTTTCGGTG CACATATTCT ATGTGGACGT
2901  AAGGCGGCAC AACCCCAGTG CCACGTTGTG AGTTGGATAG TTGTGGAAAG
      TTCCGCCGTG TTGGGGTCAC GGTGCAACAC TCAACCTATC AACACCTTTC
2951  AGTCAAATGG CTCTCCTCAA GCGTATTCAA CAAGGGGCTG AAGGATGCCC
      TCAGTTTACC GAGAGGAGTT CGCATAAGTT GTTCCCCGAC TTCCTACGGG
3001  AGAAGGTACC CCATTGTATG GGATCTGATC TGGGGCCTCG GTGCACATGC
      TCTTCCATGG GGTAACATAC CCTAGACTAG ACCCCGGAGC CACGTGTACG
3051  TTTACATGTG TTTAGTCGAG GTTAAAAAAA CGTCTAGGCC CCCCGAACCA
      AAATGTACAC AAATCAGCTC CAATTTTTTT GCAGATCCGG GGGGCTTGGT
3101  CGGGGACGTG GTTTTCCTTT GAAAAACACG ATGATAAGCT TGCCACAACC
      GCCCCTGCAC CAAAAGGAAA CTTTTTGTGC TACTATTCGA ACGGTGTTGG
3151  CACAAGGAGA CGACCTTCCA TGACCGAGTA CAAGCCCACG GTGCGCCTCG
      GTGTTCCTCT GCTGGAAGGT ACTGGCTCAT GTTCGGGTGC CACGCGGAGC
3201  CCACCCGCGA CGACGTCCCC CGGGCCGTAC GCACCCTCGC CGCCGCGTTC
      GGTGGGCGCT GCTGCAGGGG GCCCGGCATG CGTGGGAGCG GCGGCGCAAG
3251  GCCGACTACC CCGCCACGCG CCACACCGTC GACCCGGACC GCCACATCGA
      CGGCTGATGG GGCGGTGCGC GGTGTGGCAG CTGGGCCTGG CGGTGTAGCT
3301  GCGGGTCACC GAGCTGCAAG AACTCTTCCT CACGCGCGTC GGGCTCGACA
      CGCCCAGTGG CTCGACGTTC TTGAGAAGGA GTGCGCGCAG CCCGAGCTGT
3351  TCGGCAAGGT GTGGGTCGCG GACGACGGCG CCGCGGTGGC GGTCTGGACC
      AGCCGTTCCA CACCCAGCGC CTGCTGCCGC GGCGCCACCG CCAGACCTGG
3401  ACGCCGGAGA GCGTCGAAGC GGGGGCGGTG TTCGCCGAGA TCGGCCCGCG
      TGCGGCCTCT CGCAGCTTCG CCCCCGCCAC AAGCGGCTCT AGCCGGGCGC
3451  CATGGCCGAG TTGAGCGGTT CCCGGCTGGC CGCGCAGCAA CAGATGGAAG
      GTACCGGCTC AACTCGCCAA GGGCCGACCG GCGCGTCGTT GTCTACCTTC
3501  GCCTCCTGGC GCCGCACCGG CCCAAGGAGC CCGCGTGGTT CCTGGCCACC
      CGGAGGACCG CGGCGTGGCC GGGTTCCTCG GGCGCACCAA GGACCGGTGG
3551  GTCGGCGTCT CGCCCGACCA CCAGGGCAAG GGTCTGGGCA GCGCCGTCGT
      CAGCCGCAGA GCGGGCTGGT GGTCCCGTTC CCAGACCCGT CGCGGCAGCA
3601  GCTCCCCGGA GTGGAGGCGG CCGAGCGCGC CGGGGTGCCC GCCTTCCTGG
      CGAGGGGCCT CACCTCCGCC GGCTCGCGCG GCCCCACGGG CGGAAGGACC
3651  AGACCTCCGC GCCCCGCAAC CTCCCCTTCT ACGAGCGGCT CGGCTTCACC
      TCTGGAGGCG CGGGGCGTTG GAGGGGAAGA TGCTCGCCGA GCCGAAGTGG
3701  GTCACCGCCG ACGTCGAGTG CCCGAAGGAC CGCGCGACCT GGTGCATGAC
      CAGTGGCGGC TGCAGCTCAC GGGCTTCCTG GCGCGCTGGA CCACGTACTG
3751  CCGCAAGCCC GGTGCCTGAC GCCCGCCCCA CGACCCGCAG CGCCCGACCG
      GGCGTTCGGG CCACGGACTG CGGGCGGGGT GCTGGGCGTC GCGGGCTGGC
3801  AAAGGAGCGC ACGACCCCAT GGCTCCGACC GAAGCCGACC CGGGCGGCCC
      TTTCCTCGCG TGCTGGGGTA CCGAGGCTGG CTTCGGCTGG GCCCGCCGGG
3851  CGCCGACCCC GCACCCGCCC CCGAGGCCCA CCGACTCTAG ACTCGAGATC
      GCGGCTGGGG CGTGGGCGGG GGCTCCGGGT GGCTGAGATC TGAGCTCTAG
```

FIGURE 9-4

4

```
3901  GATAATCAAC CTCTGGATTA CAAAATTTGT GAAAGATTGA CTGGTATTCT
      CTATTAGTTG GAGACCTAAT GTTTTAAACA CTTTCTAACT GACCATAAGA
3951  TAACTATGTT GCTCCTTTTA CGCTATGTGG ATACGCTGCT TTAATGCCTT
      ATTGATACAA CGAGGAAAAT GCGATACACC TATGCGACGA AATTACGGAA
4001  TGTATCATGC TATTGCTTCC CGTATGGCTT TCATTTTCTC CTCCTTGTAT
      ACATAGTACG ATAACGAAGG GCATACCGAA AGTAAAAGAG GAGGAACATA
4051  AAATCCTGGT TGCTGTCTCT TTATGAGGAG TTGTGGCCCG TTGTCAGGCA
      TTTAGGACCA ACGACAGAGA AATACTCCTC AACACCGGGC AACAGTCCGT
4101  ACGTGGCGTG GTGTGCACTG TGTTTGCTGA CGCAACCCCC ACTGGTTGGG
      TGCACCGCAC CACACGTGAC ACAAACGACT GCGTTGGGGG TGACCAACCC
4151  GCATTGCCAC CACCTGTCAG CTCCTTTCCG GGACTTTCGC TTTCCCCCTC
      CGTAACGGTG GTGGACAGTC GAGGAAAGGC CCTGAAAGCG AAAGGGGGAG
4201  CCTATTGCCA CGGCGGAACT CATCGCCGCC TGCCTTGCCC GCTGCTGGAC
      GGATAACGGT GCCGCCTTGA GTAGCGGCGG ACGGAACGGG CGACGACCTG
4251  AGGGGCTCGG CTGTTGGGCA CTGACAATTC CGTGGTGTTG TCGGGGAAAT
      TCCCCGAGCC GACAACCCGT GACTGTTAAG GCACCACAAC AGCCCCTTTA
4301  CATCGTCCTT TCCTTGGCTG CTCGCCTGTG TTGCCACCTG GATTCTGCGC
      GTAGCAGGAA AGGAACCGAC GAGCGGACAC AACGGTGGAC CTAAGACGCG
4351  GGGACGTCCT TCTGCTACGT CCCTTCGGCC CTCAATCCAG CGGACCTTCC
      CCCTGCAGGA AGACGATGCA GGGAAGCCGG GAGTTAGGTC GCCTGGAAGG
4401  TTCCCGCGGC CTGCTGCCGG CTCTGCGGCC TCTTCCGCGT CTTCGCCTTC
      AAGGGCGCCG GACGACGGCC GAGACGCCGG AGAAGGCGCA GAAGCGGAAG
4451  GCCCTCAGAC GAGTCGGATC TCCCTTTGGG CCGCCTCCCC GCCTGATCGA
      CGGGAGTCTG CTCAGCCTAG AGGGAAACCC GGCGGAGGGG CGGACTAGCT
4501  TCTAGAGCTC GCTGATCAGC CTCGACTGTG CCTTCTAGTT GCCAGCCATC
      AGATCTCGAG CGACTAGTCG GAGCTGACAC GGAAGATCAA CGGTCGGTAG
4551  TGTTGTTTGC CCCTCCCCCG TGCCTTCCTT GACCCTGGAA GGTGCCACTC
      ACAACAAACG GGGAGGGGGC ACGGAAGGAA CTGGGACCTT CCACGGTGAG
4601  CCACTGTCCT TTCCTAATAA AATGAGGAAA TTGCATCGCA TTGTCTGAGT
      GGTGACAGGA AAGGATTATT TTACTCCTTT AACGTAGCGT AACAGACTCA
4651  AGGTGTCATT CTATTCTGGG GGGTGGGGTG GGGCAGGACA GCAAGGGGGA
      TCCACAGTAA GATAAGACCC CCCACCCCAC CCCGTCCTGT CGTTCCCCCT
4701  GGATTGGGAA GACAATAGCA GGCATGCTGG GGATGCGGTG GGCTCTATGG
      CCTAACCCTT CTGTTATCGT CCGTACGACC CCTACGCCAC CCGAGATACC
4751  CTTCTGAGGC GGAAAGAACC AGCTGGGGCT CGAGTGCATT CTAGTTGTGG
      GAAGACTCCG CCTTTCTTGG TCGACCCCGA GCTCACGTAA GATCAACACC
4801  TTTGTCCAAA CTCATCAATG TATCTTATCA TGTCTGTATA CCGTCGACCT
      AAACAGGTTT GAGTAGTTAC ATAGAATAGT ACAGACATAT GGCAGCTGGA
4851  CTAGCTAGAG CTTGGCGTAA TCATGGTCAT AGCTGTTTCC TGTGTGAAAT
      GATCGATCTC GAACCGCATT AGTACCAGTA TCGACAAAGG ACACACTTTA
4901  TGTTATCCGC TCACAATTCC ACACAACATA CGAGCCGGAA GCATAAAGTG
      ACAATAGGCG AGTGTTAAGG TGTGTTGTAT GCTCGGCCTT CGTATTTCAC
4951  TAAAGCCTGG GGTGCCTAAT GAGTGAGCTA ACTCACATTA ATTGCGTTGC
      ATTTCGGACC CCACGGATTA CTCACTCGAT TGAGTGTAAT TAACGCAACG
5001  GCTCACTGCC CGCTTTCCAG TCGGGAAACC TGTCGTGCCA GCTGCATTAA
      CGAGTGACGG GCGAAAGGTC AGCCCTTTGG ACAGCACGGT CGACGTAATT
```

FIGURE 9-5

```
5051  TGAATCGGCC  AACGCGCGGG  GAGAGGCGGT  TTGCGTATTG  GGCGCTCTTC
      ACTTAGCCGG  TTGCGCGCCC  CTCTCCGCCA  AACGCATAAC  CCGCGAGAAG
5101  CGCTTCCTCG  CTCACTGACT  CGCTGCGCTC  GGTCGTTCGG  CTGCGGCGAG
      GCGAAGGAGC  GAGTGACTGA  GCGACGCGAG  CCAGCAAGCC  GACGCCGCTC
5151  CGGTATCAGC  TCACTCAAAG  GCGGTAATAC  GGTTATCCAC  AGAATCAGGG
      GCCATAGTCG  AGTGAGTTTC  CGCCATTATG  CCAATAGGTG  TCTTAGTCCC
5201  GATAACGCAG  GAAAGAACAT  GTGAGCAAAA  GGCCAGCAAA  AGGCCAGGAA
      CTATTGCGTC  CTTTCTTGTA  CACTCGTTTT  CCGGTCGTTT  TCCGGTCCTT
5251  CCGTAAAAAG  GCCGCGTTGC  TGGCGTTTTT  CCATAGGCTC  CGCCCCCCTG
      GGCATTTTTC  CGGCGCAACG  ACCGCAAAAA  GGTATCCGAG  GCGGGGGGAC
5301  ACGAGCATCA  CAAAAATCGA  CGCTCAAGTC  AGAGGTGGCG  AAACCCGACA
      TGCTCGTAGT  GTTTTTAGCT  GCGAGTTCAG  TCTCCACCGC  TTTGGGCTGT
5351  GGACTATAAA  GATACCAGGC  GTTTCCCCCT  GGAAGCTCCC  TCGTGCGCTC
      CCTGATATTT  CTATGGTCCG  CAAAGGGGGA  CCTTCGAGGG  AGCACGCGAG
5401  TCCTGTTCCG  ACCCTGCCGC  TTACCGGATA  CCTGTCCGCC  TTTCTCCCTT
      AGGACAAGGC  TGGGACGGCG  AATGGCCTAT  GGACAGGCGG  AAAGAGGGAA
5451  CGGGAAGCGT  GGCGCTTTCT  CAATGCTCAC  GCTGTAGGTA  TCTCAGTTCG
      GCCCTTCGCA  CCGCGAAAGA  GTTACGAGTG  CGACATCCAT  AGAGTCAAGC
5501  GTGTAGGTCG  TTCGCTCCAA  GCTGGGCTGT  GTGCACGAAC  CCCCCGTTCA
      CACATCCAGC  AAGCGAGGTT  CGACCCGACA  CACGTGCTTG  GGGGGCAAGT
5551  GCCCGACCGC  TGCGCCTTAT  CCGGTAACTA  TCGTCTTGAG  TCCAACCCGG
      CGGGCTGGCG  ACGCGGAATA  GGCCATTGAT  AGCAGAACTC  AGGTTGGGCC
5601  TAAGACACGA  CTTATCGCCA  CTGGCAGCAG  CCACTGGTAA  CAGGATTAGC
      ATTCTGTGCT  GAATAGCGGT  GACCGTCGTC  GGTGACCATT  GTCCTAATCG
5651  AGAGCGAGGT  ATGTAGGCGG  TGCTACAGAG  TTCTTGAAGT  GGTGGCCTAA
      TCTCGCTCCA  TACATCCGCC  ACGATGTCTC  AAGAACTTCA  CCACCGGATT
5701  CTACGGCTAC  ACTAGAAGGA  CAGTATTTGG  TATCTGCGCT  CTGCTGAAGC
      GATGCCGATG  TGATCTTCCT  GTCATAAACC  ATAGACGCGA  GACGACTTCG
5751  CAGTTACCTT  CGGAAAAAGA  GTTGGTAGCT  CTTGATCCGG  CAAACAAACC
      GTCAATGGAA  GCCTTTTTCT  CAACCATCGA  GAACTAGGCC  GTTTGTTTGG
5801  ACCGCTGGTA  GCGGTGGTTT  TTTTGTTTGC  AAGCAGCAGA  TTACGCGCAG
      TGGCGACCAT  CGCCACCAAA  AAAACAAACG  TTCGTCGTCT  AATGCGCGTC
5851  AAAAAAAGGA  TCTCAAGAAG  ATCCTTTGAT  CTTTTCTACG  GGGTCTGACG
      TTTTTTTCCT  AGAGTTCTTC  TAGGAAACTA  GAAAAGATGC  CCCAGACTGC
5901  CTCAGTGGAA  CGAAAACTCA  CGTTAAGGGA  TTTTGGTCAT  GAGATTATCA
      GAGTCACCTT  GCTTTTGAGT  GCAATTCCCT  AAAACCAGTA  CTCTAATAGT
5951  AAAAGGATCT  TCACCTAGAT  CCTTTTAAAT  TAAAAATGAA  GTTTTAAATC
      TTTTCCTAGA  AGTGGATCTA  GGAAAATTTA  ATTTTTACTT  CAAAATTTAG
6001  AATCTAAAGT  ATATATGAGT  AAACTTGGTC  TGACAGTTAC  CAATGCTTAA
      TTAGATTTCA  TATATACTCA  TTTGAACCAG  ACTGTCAATG  GTTACGAATT
6051  TCAGTGAGGC  ACCTATCTCA  GCGATCTGTC  TATTTCGTTC  ATCCATAGTT
      AGTCACTCCG  TGGATAGAGT  CGCTAGACAG  ATAAAGCAAG  TAGGTATCAA
6101  GCCTGACTCC  CCGTCGTGTA  GATAACTACG  ATACGGGAGG  GCTTACCATC
      CGGACTGAGG  GGCAGCACAT  CTATTGATGC  TATGCCCTCC  CGAATGGTAG
6151  TGGCCCCAGT  GCTGCAATGA  TACCGCGAGA  CCCACGCTCA  CCGGCTCCAG
      ACCGGGGTCA  CGACGTTACT  ATGGCGCTCT  GGGTGCGAGT  GGCCGAGGTC
```

FIGURE 9-6

6

```
6201  ATTTATCAGC AATAAACCAG CCAGCCGGAA GGGCCGAGCG CAGAAGTGGT
      TAAATAGTCG TTATTTGGTC GGTCGGCCTT CCCGGCTCGC GTCTTCACCA
6251  CCTGCAACTT TATCCGCCTC CATCCAGTCT ATTAATTGTT GCCGGGAAGC
      GGACGTTGAA ATAGGCGGAG GTAGGTCAGA TAATTAACAA CGGCCCTTCG
6301  TAGAGTAAGT AGTTCGCCAG TTAATAGTTT GCGCAACGTT GTTGCCATTG
      ATCTCATTCA TCAAGCGGTC AATTATCAAA CGCGTTGCAA CAACGGTAAC
6351  CTACAGGCAT CGTGGTGTCA CGCTCGTCGT TTGGTATGGC TTCATTCAGC
      GATGTCCGTA GCACCACAGT GCGAGCAGCA AACCATACCG AAGTAAGTCG
6401  TCCGGTTCCC AACGATCAAG GCGAGTTACA TGATCCCCCA TGTTGTGCAA
      AGGCCAAGGG TTGCTAGTTC CGCTCAATGT ACTAGGGGGT ACAACACGTT
6451  AAAAGCGGTT AGCTCCTTCG GTCCTCCGAT CGTTGTCAGA AGTAAGTTGG
      TTTTCGCCAA TCGAGGAAGC CAGGAGGCTA GCAACAGTCT TCATTCAACC
6501  CCGCAGTGTT ATCACTCATG GTTATGGCAG CACTGCATAA TTCTCTTACT
      GGCGTCACAA TAGTGAGTAC CAATACCGTC GTGACGTATT AAGAGAATGA
6551  GTCATGCCAT CCGTAAGATG CTTTTCTGTG ACTGGTGAGT ACTCAACCAA
      CAGTACGGTA GGCATTCTAC GAAAAGACAC TGACCACTCA TGAGTTGGTT
6601  GTCATTCTGA GAATAGTGTA TGCGGCGACC GAGTTGCTCT TGCCCGGCGT
      CAGTAAGACT CTTATCACAT ACGCCGCTGG CTCAACGAGA ACGGGCCGCA
6651  CAATACGGGA TAATACCGCG CCACATAGCA GAACTTTAAA AGTGCTCATC
      GTTATGCCCT ATTATGGCGC GGTGTATCGT CTTGAAATTT TCACGAGTAG
6701  ATTGGAAAAC GTTCTTCGGG GCGAAAACTC TCAAGGATCT TACCGCTGTT
      TAACCTTTTG CAAGAAGCCC CGCTTTTGAG AGTTCCTAGA ATGGCGACAA
6751  GAGATCCAGT TCGATGTAAC CCACTCGTGC ACCCAACTGA TCTTCAGCAT
      CTCTAGGTCA AGCTACATTG GGTGAGCACG TGGGTTGACT AGAAGTCGTA
6801  CTTTTACTTT CACCAGCGTT TCTGGGTGAG CAAAAACAGG AAGGCAAAAT
      GAAAATGAAA GTGGTCGCAA AGACCCACTC GTTTTTGTCC TTCCGTTTTA
6851  GCCGCAAAAA AGGGAATAAG GGCGACACGG AAATGTTGAA TACTCATACT
      CGGCGTTTTT TCCCTTATTC CCGCTGTGCC TTTACAACTT ATGAGTATGA
6901  CTTCCTTTTT CAATATTATT GAAGCATTTA TCAGGGTTAT TGTCTCATGA
      GAAGGAAAAA GTTATAATAA CTTCGTAAAT AGTCCCAATA ACAGAGTACT
6951  GCGGATACAT ATTTGAATGT ATTTAGAAAA ATAAACAAAT AGGGGTTCCG
      CGCCTATGTA TAAACTTACA TAAATCTTTT TATTTGTTTA TCCCCAAGGC
7001  CGCACATTTC CCCGAAAAGT GCCACCTGAC GTC
      GCGTGTAAAG GGGCTTTTCA CGGTGGACTG CAG
```

FIGURE 9-7

FIG. 10 (1-12)

pCEP4W/hepEK

```
  1   TCGAGCGGCC GCTTTAAACT CAATGGTGAT GGTGATGATG ACCGGTACGC
      AGCTCGCCGG CGAAATTTGA GTTACCACTA CCACTACTAC TGGCCATGCG
 -3                        H    H    H    H    H    H   G  T  R  T

 51   GTAGAATCGA GACCGAGGAG AGGGTTAGGG ATAGGCTTAC CGAATTCGAG
      CATCTTAGCT CTGGCTCCTC TCCCAATCCC TATCCGAATG GCTTAAGCTC
 -3   ·T  S  D  L    G    L    L    P    N  P   I   P  K  G  F  E  L

101   CTGGGTCACC ATGCCGCTGG CTTCGGAGTG AGTCTTTATG GCCTGGAAGA
      GACCCAGTGG TACGGCGACC GAAGCCTCAC TCAGAAATAC CGGACCTTCT
 -3    Q  T  V  M    G    S    A    E    S  H   T   K  I  A  Q  F  I·

151   TCCACTCCCG GAAGTCACTG ACTTTGGTGT AGACGCCTGG CTTCTGGGCC
      AGGTGAGGGC CTTCAGTGAC TGAAACCACA TCTGCGGACC GAAGACCCGG
 -3   ·I  W  E  R    F    D    S    V    K  T   Y   V  G  P  K  Q  A  L

201   AGGGCACAGC CAGTGCCCCA ACTCACAATG CCACACAGCC GCCAACGTGG
      TCCCGTGTCG GTCACGGGGT TGAGTGTTAC GGTGTGTCGG CGGTTGCACC
 -3   ·L  A  C  G    T    G    W    S    V  I   G   C  L  R  W  R  P

251   CGTCCGAGAG ATGCTGTCCT CACACACAAA GGGACCACCG CTGTCGCCCT
      GCAGGCTCTC TACGACAGGA GTGTGTGTTT CCCTGGTGGC GACAGCGGGA
 -3    T  R  S  I    S    D    E    C    V  F   P   G  G  S  D  G  Q·

301   GGCAGGCATC AATGCCACCC TCGGGGTAGC CAGCACAGAA CATCTTGGGC
      CCGTCCGTAG TTACGGTGGG AGCCCCATCG GTCGTGTCTT GTAGAACCCG
 -3   ·Q  C  A  D    I    G    G    E    P  Y   G   A  C  F  M  K  P  K

351   TTGATCTGGT TTCCATAGAA ATCAGCGCCA TTGCAGACAT CATTGCTGAT
      AACTAGACCA AAGGTATCTT TAGTCGCGGT AACGTCTGTA GTAACGACTA
 -3   ·K  I  Q  N    G    Y    F    D    A  G   N   C  V  D  N  S  I

401   TATGGGGACT CGAGCCTCCT GGAGTACCCC GGCCTGTTGG CCATAGTACT
      ATACCCCTGA GCTCGGAGGA CCTCATGGGG CCGGACAACC GGTATCATGA
 -3    I  P  V  R    A    E    Q    L    V  G   A   Q  Q  G  Y  Y  Q·

451   GCGTGTTGCC CCAGCCCGTC ACGGTACAGA TCTTGCCATC CACCAGGGCC
      CGCACAACGG GGTCGGGCAG TGCCATGTCT AGAACGGTAG GTGGTCCCGG
 -3   ·Q  T  N  G    W    G    T    V    T  C   I   K  G  D  V  L  A  Q

501   TGGCCGGCAG CTGGGAGGCA CACAGGCTGG ATGTATTCTG TGAGGGGCAG
      ACCGGCCGTC GACCCTCCGT GTGTCCGACC TACATAAGAC ACTCCCCGTC
 -3   ·Q  G  A  A    P    L    C    V    P  Q   I   Y  E  T  L  P  L

551   GGGACTGGAG AGGTGGACCA GGGCAATATC GTTGCTGTTC TCCTCGCTGT
      CCCTGACCTC TCCACCTGGT CCCGTTATAG CAACGACAAG AGGAGCGACA
 -3    P  S  S  L    H    V    L    A    I  D   N   S  N  E  E  S  N·

601   TGGGGTCCCG AAAGGGAAGA TAGCCCCCGT GGTAGACCAC AGCCTGCACC
      ACCCCAGGGC TTTCCCTTCT ATCGGGGGCA CCATCTGGTG TCGGACGTGG
 -3   ·N  P  D  R    F    P    L    Y    G  G   H   Y  V  V  A  Q  V  G

651   CCCAGCTGCA GACCGTGGGG AGAGGCCTGG CCACGGCAC CGGCAAACAC
      GGGTCGACGT CTGGCACCCC TCTCCGGACC CGGTGCCGTG GCCGTTTGTG
 -3   ·G  L  Q  L    G    H    P    S    A  Q   A   V  A  G  A  F  V

701   TCGCCATCGG GACAGGACCC GGTTCCGCTC CGGGAAGCAG TGGGCGGCTG
      AGCGGTAGCC CTGTCCTGGG CCAAGGCGAG GCCCTTCGTC ACCCGCCGAC
 -3    R  W  R  S    L    V    R    N    R  E   P   F  C  H  A  A  T·
```

FIGURE 10-1

Fig 10

1

```
 751  TCAGCACCCA GTCCCCGGAG AGCAGGGATC CCCCACAGAG GTGTGCTCCA
      AGTCGTGGGT CAGGGGCCTC TCGTCCCTAG GGGGTGTCTC CACACGAGGT
  -3  ·T  L   V   D   G   S   L   L   S   G   G   C   L   H   A   G   D
 801  TCATAGCGAA GGCTGACTTG CCACGGCCAC CGGCCCAAGC TGGTGTCCCG
      AGTATCGCTT CCGACTGAAC GGTGCCGGTG GCCGGGTTCG ACCACAGGGC
  -3  ·D  Y   R   L   S   V   Q   W   P   W   R   G   L   S   T   D   R
 851  GCCTCCCACG ATCTTGTCGT CGTCGTCCAC GGGCAGCTTC CTGCGGCCAC
      CGGAGGGTGC TAGAACAGCA GCAGCAGGTG CCCGTCGAAG GACGCCGGTG
  -3   G  G   V   I   K   D   D   D   D   V   P   L   K   R   R   G  C·
 901  AGTCTTGGCA GATGGCGGCC AAGAAACGGC CTCTGGGGCA ATCACACACG
      TCAGAACCGT CTACCGCCGG TTCTTTGCCG GAGACCCCGT TAGTGTGTGC
  -3  ·C  D   Q   C   I   A   A   L   F   R   G   R   P   C   D   C   V   S
 951  GAGATGACCT CCAGCAGCCT CTGGGTGTGG GGCAGCCTCC CCTCGTCCAC
      CTCTACTGGA GGTCGTCGGA GACCCACACC CCGTCGGAGG GGAGCAGGTG
  -3  ·S  I   V   E   L   L   R   Q   T   H   P   L   R   G   E   D   V
1001  ACAGAAGAAG CCCGACGTGC CATTGGCGCC CGCCGTTCGC ACGTCCAGCT
      TGTCTTCTTC GGGCTGCACG GTAACCGCGG GCGGCAAGCG TGCAGGTCGA
  -3   C  F   F   G   S   T   G   N   A   G   A   T   R   V   D   L  E·
1051  CGGAGTGGGT CAGTGCCCTG AGGAAGCCCA TCTCCTCGCA GCTGAGTCCG
      GCCTCACCCA GTCACGGGAC TCCTTCGGGT AGAGGAGCGT CGACTCAGGC
  -3  ·E  S   H   T   L   A   R   L   F   G   M   E   E   C   S   L   G  A·
1101  GCTACCCTGG CGTTAGAGCG CGAGGAGCAC AGCAGCCGCC ACGTCCCTTC
      CGATGGGACC GCAATCTCGC GCTCCTCGTG TCGTCGGCGG TGCAGGGAAG
  -3  ·A  V   R   A   N   S   R   S   S   C   L   L   R   W   T   G   E
1151  CGTCTTGTCA AAGACCATGA GCCGAGCGTC CGCAGAGCTG ACCTGCACTG
      GCAGAACAGT TTCTGGTACT CGGCTCGCAG GCGTCTCGAC TGGACGTGAC
  -3   T  K   D   F   V   M   L   R   A   D   A   S   S   V   Q   V  P·
1201  GGTACAGCGG CTCCTGGTCA CTCCTATCCG GAGCGTCACC AGTGGAACCT
      CCATGTCGCC GAGGACCAGT GAGGATAGGC CTCGCAGTGG TCACCTTGGA
  -3  ·P  Y   L   P   E   Q   D   S   R   D   P   A   D   G   T   S   G  P·
1251  GGAACCCAGA GCAGCAGTAC CCATAGCAGG AGTGTGTCTG TCTCCATGGT
      CCTTGGGTCT CGTCGTCATG GGTATCGTCC TCACACAGAC AGAGGTACCA
  -3  ·P  V   W   L   L   L   V   W   L   L   L   T   D   T   E   M
1301  GGCGATCTGG TACCCAGCTT CTAGAGATCT GACGGTTCAC TAAACGAGCT
      CCGCTAGACC ATGGGTCGAA GATCTCTAGA CTGCCAAGTG ATTTGCTCGA
1351  CTGCTTATAT AGACCTCCCA CCGTACACGC CTACCGCCCA TTTGCGTCAA
      GACGAATATA TCTGGAGGGT GGCATGTGCG GATGGCGGGT AAACGCAGTT
1401  CGGGGCGGGG TTATTACGAC ATTTTGGAAA GTCCCGTTGA TTTTGGTGCC
      GCCCCGCCCC AATAATGCTG TAAAACCTTT CAGGGCAACT AAAACCACGG
1451  AAAACAAACT CCCATTGACG TCAATGGGGT GGAGACTTGG AAATCCCCGT
      TTTTGTTTGA GGGTAACTGC AGTTACCCCA CCTCTGAACC TTTAGGGGCA
1501  GAGTCAAACC GCTATCCACG CCCATTGGTG TACTGCCAAA ACCGCATCAC
      CTCAGTTTGG CGATAGGTGC GGGTAACCAC ATGACGGTTT TGGCGTAGTG
1551  CATGGTAATA GCGATGACTA ATACGTAGAT GTACTGCCAA GTAGGAAAGT
      GTACCATTAT CGCTACTGAT TATGCATCTA CATGACGGTT CATCCTTTCA
```

FIGURE 10-2

2

```
1601  CCCGTAAGGT CATGTACTGG GCATAATGCC AGGCGGGCCA TTTACCGTCA
      GGGCATTCCA GTACATGACC CGTATTACGG TCCGCCCGGT AAATGGCAGT
1651  TTGACGTCAA TAGGGGGCGG ACTTGGCATA TGATACACTT GATGTACTGC
      AACTGCAGTT ATCCCCCGCC TGAACCGTAT ACTATGTGAA CTACATGACG
1701  CAAGTGGGCA GTTTACCGTA AATACTCCAC CCATTGACGT CAATGGAAAG
      GTTCACCCGT CAAATGGCAT TTATGAGGTG GGTAACTGCA GTTACCTTTC
1751  TCCCTATTGG CGTTACTATG GAACATACG TCATTATTGA CGTCAATGGG
      AGGGATAACC GCAATGATAC CCTTGTATGC AGTAATAACT GCAGTTACCC
1801  CGGGGGTCGT TGGGCGGTCA GCCAGGCGGG CCATTTACCG TAAGTTATGT
      GCCCCCAGCA ACCCGCCAGT CGGTCCGCCC GGTAAATGGC ATTCAATACA
1851  AACGCGGAAC TCCATATATG GGCTATGAAC TAATGACCCC GTAATTGATT
      TTGCGCCTTG AGGTATATAC CCGATACTTG ATTACTGGGG CATTAACTAA
1901  ACTATTAATA ACTAGTCAAT AATCAATGTC AACATGGCGG TCATATTGGA
      TGATAATTAT TGATCAGTTA TTAGTTACAG TTGTACCGCC AGTATAACCT
1951  CATGAGCCAA TATAAATGTA CATATTATGA TATAGATACA ACGTATGCAA
      GTACTCGGTT ATATTTACAT GTATAATACT ATATCTATGT TGCATACGTT
2001  TGGCCAATAG CCAATATTGA TTTATGCTAT ATAACCAATG ACTAATATGG
      ACCGGTTATC GGTTATAACT AAATACGATA TATTGGTTAC TGATTATACC
2051  CTAATTGCCA ATATTGATTC AATGTATAGA TCTTCCATAC CTACCAGTTC
      GATTAACGGT TATAACTAAG TTACATATCT AGAAGGTATG GATGGTCAAG
2101  TGCGCCTGCA GCAATGCAAC AACGTTGCCC GGATCTGCGA TGATAAGCTG
      ACGCGGACGT CGTTACGTTG TTGCAACGGG CCTAGACGCT ACTATTCGAC
2151  TCAAACATGA GAATTGGTCG ACTAGCTTGG CACGCCAGAA ATCCGCGCGG
      AGTTTGTACT CTTAACCAGC TGATCGAACC GTGCGGTCTT TAGGCGCGCC
2201  TGGTTTTTGG GGGTCGGGGG TGTTTGGCAG CCACAGACGC CCGGTGTTCG
      ACCAAAAACC CCCAGCCCCC ACAAACCGTC GGTGTCTGCG GGCCACAAGC
2251  TGTCGCGCCA GTACATGCGG TCCATGCCCA GGCCATCCAA AAACCATGGG
      ACAGCGCGGT CATGTACGCC AGGTACGGGT CCGGTAGGTT TTTGGTACCC
2301  TCTGTCTGCT CAGTCCAGTC GTGGACCAGA CCCCACGCAA CGCCCAAAAT
      AGACAGACGA GTCAGGTCAG CACCTGGTCT GGGGTGCGTT GCGGGTTTTA
2351  AATAACCCCC ACGAACCATA AACCATTCCC CATGGGGGAC CCCGTCCCTA
      TTATTGGGGG TGCTTGGTAT TTGGTAAGGG GTACCCCCTG GGGCAGGGAT
2401  ACCCACGGGG CCAGTGGCTA TGGCAGGGCC TGCCGCCCCG ACGTTGGCTG
      TGGGTGCCCC GGTCACCGAT ACCGTCCCGG ACGGCGGGGC TGCAACCGAC
2451  CGAGCCCTGG GCCTTCACCC GAACTTGGGG GGTGGGGTGG GGAAAAGGAA
      GCTCGGGACC CGGAAGTGGG CTTGAACCCC CCACCCCACC CCTTTTCCTT
2501  GAAACGCGGG CGTATTGGCC CCAATGGGGT CTCGGTGGGG TATCGACAGA
      CTTTGCGCCC GCATAACCGG GGTTACCCCA GAGCCACCCC ATAGCTGTCT
2551  GTGCCAGCCC TGGGACCGAA CCCCGCGTTT ATGAACAAAC GACCCAACAC
      CACGGTCGGG ACCCTGGCTT GGGGCGCAAA TACTTGTTTG CTGGGTTGTG
2601  CCGTGCGTTT TATTCTGTCT TTTTATTGCC GTCATAGCGC GGGTTCCTTC
      GGCACGCAAA ATAAGACAGA AAAATAACGG CAGTATCGCG CCCAAGGAAG
2651  CGGTATTGTC TCCTTCCGTG TTTCAGTTAG CCTCCCCCAT CTCCCCTATT
      GCCATAACAG AGGAAGGCAC AAAGTCAATC GGAGGGGGTA GAGGGGATAA
2701  CCTTTGCCCT CGGACGAGTG CTGGGGCGTC GGTTTCCACT ATCGGCGAGT
      GGAAACGGGA GCCTGCTCAC GACCCCGCAG CCAAAGGTGA TAGCCGCTCA
```

FIGURE 10-3

3

```
2751  ACTTCTACAC AGCCATCGGT CCAGACGGCC GCGCTTCTGC GGGCGATTTG
      TGAAGATGTG TCGGTAGCCA GGTCTGCCGG CGCGAAGACG CCCGCTAAAC
2801  TGTACGCCCG ACAGTCCCGG CTCCGGATCG GACGATTGCG TCGCATCGAC
      ACATGCGGGC TGTCAGGGCC GAGGCCTAGC CTGCTAACGC AGCGTAGCTG
2851  CCTGCGCCCA AGCTGCATCA TCGAAATTGC CGTCAACCAA GCTCTGATAG
      GGACGCGGGT TCGACGTAGT AGCTTTAACG GCAGTTGGTT CGAGACTATC
2901  AGTTGGTCAA GACCAATGCG GAGCATATAC GCCCGGAGCC GCGGCGATCC
      TCAACCAGTT CTGGTTACGC CTCGTATATG CGGGCCTCGG CGCCGCTAGG
2951  TGCAAGCTCC GGATGCCTCC GCTCGAAGTA GCGCGTCTGC TGCTCCATAC
      ACGTTCGAGG CCTACGGAGG CGAGCTTCAT CGCGCAGACG ACGAGGTATG
3001  AAGCCAACCA CGGCCTCCAG AAGAAGATGT TGGCGACCTC GTATTGGGAA
      TTCGGTTGGT GCCGGAGGTC TTCTTCTACA ACCGCTGGAG CATAACCCTT
3051  TCCCCGAACA TCGCCTCGCT CCAGTCAATG ACCGCTGTTA TGCGGCCATT
      AGGGGCTTGT AGCGGAGCGA GGTCAGTTAC TGGCGACAAT ACGCCGGTAA
3101  GTCCGTCAGG ACATTGTTGG AGCCGAAATC CGCGTGCACG AGGTGCCGGA
      CAGGCAGTCC TGTAACAACC TCGGCTTTAG GCGCACGTGC TCCACGGCCT
3151  CTTCGGGGCA GTCCTCGGCC CAAAGCATCA GCTCATCGAG AGCCTGCGCG
      GAAGCCCCGT CAGGAGCCGG GTTTCGTAGT CGAGTAGCTC TCGGACGCGC
3201  ACGGACGCAC TGACGGTGTC GTCCATCACA GTTTGCCAGT GATACACATG
      TGCCTGCGTG ACTGCCACAG CAGGTAGTGT CAAACGGTCA CTATGTGTAC
3251  GGGATCAGCA ATCGCGCATA TGAAATCACG CCATGTAGTG TATTGACCGA
      CCCTAGTCGT TAGCGCGTAT ACTTTAGTGC GGTACATCAC ATAACTGGCT
3301  TTCCTTGCGG TCCGAATGGG CCGAACCCGC TCGTCTGGCT AAGATCGGCC
      AAGGAACGCC AGGCTTACCC GGCTTGGGCG AGCAGACCGA TTCTAGCCGG
3351  GCAGCGATCG CATCCATGGC CTCCGCGACC GGCTGCAGAA CAGCGGGCAG
      CGTCGCTAGC GTAGGTACCG GAGGCGCTGG CCGACGTCTT GTCGCCCGTC
3401  TTCGGTTTCA GGCAGGTCTT GCAACGTGAC ACCCTGTGCA CGGCGGGAGA
      AAGCCAAAGT CCGTCCAGAA CGTTGCACTG TGGGACACGT GCCGCCCTCT
3451  TGCAATAGGT CAGGCTCTCG CTGAATTCCC CAATGTCAAG CACTTCCGGA
      ACGTTATCCA GTCCGAGAGC GACTTAAGGG GTTACAGTTC GTGAAGGCCT
3501  ATCGGGAGCG CGGCCGATGC AAAGTGCCGA TAAACATAAC GATCTTTGTA
      TAGCCCTCGC GCCGGCTACG TTTCACGGCT ATTTGTATTG CTAGAAACAT
3551  GAAACCATCG GCGCAGCTAT TTACCCGCAG GACATATCCA CGCCCTCCTA
      CTTTGGTAGC CGCGTCGATA AATGGGCGTC CTGTATAGGT GCGGGAGGAT
3601  CATCGAAGCT GAAAGCACGA GATTCTTCGC CCTCCGAGAG CTGCATCAGG
      GTAGCTTCGA CTTTCGTGCT CTAAGAAGCG GGAGGCTCTC GACGTAGTCC
3651  TCGGAGACGC TGTCGAACTT TTCGATCAGA AACTTCTCGA CAGACGTCGC
      AGCCTCTGCG ACAGCTTGAA AAGCTAGTCT TTGAAGAGCT GTCTGCAGCG
3701  GGTGAGTTCA GGCTTTTTCA TATCTCATTG CCCGGGATCT GCGGCACGCT
      CCACTCAAGT CCGAAAAAGT ATAGAGTAAC GGGCCCTAGA CGCCGTGCGA
3751  GTTGACGCTG TTAAGCGGGT CGCTGCAGGG TCGCTCGGTG TTCGAGGCCA
      CAACTGCGAC AATTCGCCCA GCGACGTCCC AGCGAGCCAC AAGCTCCGGT
3801  CACGCGTCAC CTTAATATGC GAAGTGGACC TGGGACCGCG CCGCCCCGAC
      GTGCGCAGTG GAATTATACG CTTCACCTGG ACCCTGGCGC GGCGGGGCTG
3851  TGCATCTGCG TGTTCGAATT CGCCAATGAC AAGACGCTGG GCGGGGTTTG
      ACGTAGACGC ACAAGCTTAA GCGGTTACTG TTCTGCGACC CGCCCCAAAC
```

FIGURE 10-4

```
3901  TGTCATCATA GAACTAAAGA CATGCAAATA TATTTCTTCC GGGGACACCG
      ACAGTAGTAT CTTGATTTCT GTACGTTTAT ATAAAGAAGG CCCCTGTGGC
3951  CCAGCAAACG CGAGCAACGG GCCACGGGGA TGAAGCAGGG CATGGCGGCC
      GGTCGTTTGC GCTCGTTGCC CGGTGCCCCT ACTTCGTCCC GTACCGCCGG
4001  GACGCGCTGG GCTACGTCTT GCTGGCGTTC GCGACGCGAG GCTGGATGGC
      CTGCGCGACC CGATGCAGAA CGACCGCAAG CGCTGCGCTC CGACCTACCG
4051  CTTCCCCATT ATGATTCTTC TCGCTTCCGG CGGCATCGGG ATGCCCGCGT
      GAAGGGGTAA TACTAAGAAG AGCGAAGGCC GCCGTAGCCC TACGGGCGCA
4101  TGCAGGCCAT GCTGTCCAGG CAGGTAGATG ACGACCATCA GGGACAGCTT
      ACGTCCGGTA CGACAGGTCC GTCCATCTAC TGCTGGTAGT CCCTGTCGAA
4151  CAAGGATCGC TCGCGGCTCT TACCAGCCTA ACTTCGATCA CTGGACCGCT
      GTTCCTAGCG AGCGCCGAGA ATGGTCGGAT TGAAGCTAGT GACCTGGCGA
4201  GATCGTCACG GCGATTTATG CCGCCTCGGC GAGCACATGG AACGGGTTGG
      CTAGCAGTGC CGCTAAATAC GGCGGAGCCG CTCGTGTACC TTGCCCAACC
4251  CATGGATTGT AGGCGCCGCC CTATACCTTG TCTGCCTCCC CGCGTTGCGT
      GTACCTAACA TCCGCGGCGG GATATGGAAC AGACGGAGGG GCGCAACGCA
4301  CGCGGTGCAT GGAGCCGGGC CACCTCGACC TGAATGGAAG CCGGCGGCAC
      GCGCCACGTA CCTCGGCCCG GTGGAGCTGG ACTTACCTTC GGCCGCCGTG
4351  CTCGCTAACG GATTCACCAC TCCAAGAATT GGAGCCAATC AATTCTTGCG
      GAGCGATTGC CTAAGTGGTG AGGTTCTTAA CCTCGGTTAG TTAAGAACGC
4401  GAGAACTGTG AATGCGCAAA CCAACCCTTG GCAGAACATA TCCATCGCGT
      CTCTTGACAC TTACGCGTTT GGTTGGGAAC CGTCTTGTAT AGGTAGCGCA
4451  CCGCCATCTC CAGCAGCCGC ACGCGGCGCA GCAAAAGGCC AGGAACCGTA
      GGCGGTAGAG GTCGTCGGCG TGCGCCGCGT CGTTTTCCGG TCCTTGGCAT
4501  AAAAGGCCGC GTTGCTGGCG TTTTTCCATA GGCTCCGCCC CCCTGACGAG
      TTTTCCGGCG CAACGACCGC AAAAAGGTAT CCGAGGCGGG GGGACTGCTC
4551  CATCACAAAA ATCGACGCTC AAGTCAGAGG TGGCGAAACC CGACAGGACT
      GTAGTGTTTT TAGCTGCGAG TTCAGTCTCC ACCGCTTTGG GCTGTCCTGA
4601  ATAAAGATAC CAGGCGTTTC CCCCTGGAAG CTCCCTCGTG CGCTCTCCTG
      TATTTCTATG GTCCGCAAAG GGGGACCTTC GAGGGAGCAC GCGAGAGGAC
4651  TTCCGACCCT GCCGCTTACC GGATACCTGT CCGCCTTTCT CCCTTCGGGA
      AAGGCTGGGA CGGCGAATGG CCTATGGACA GGCGGAAAGA GGGAAGCCCT
4701  AGCGTGGCGC TTTCTCATAG CTCACGCTGT AGGTATCTCA GTTCGGTGTA
      TCGCACCGCG AAAGAGTATC GAGTGCGACA TCCATAGAGT CAAGCCACAT
4751  GGTCGTTCGC TCCAAGCTGG GCTGTGTGCA CGAACCCCCC GTTCAGCCCG
      CCAGCAAGCG AGGTTCGACC CGACACACGT GCTTGGGGGG CAAGTCGGGC
4801  ACCGCTGCGC CTTATCCGGT AACTATCGTC TTGAGTCCAA CCCGGTAAGA
      TGGCGACGCG GAATAGGCCA TTGATAGCAG AACTCAGGTT GGGCCATTCT
4851  CACGACTTAT CGCCACTGGC AGCAGCCACT GGTAACAGGA TTAGCAGAGC
      GTGCTGAATA GCGGTGACCG TCGTCGGTGA CCATTGTCCT AATCGTCTCG
4901  GAGGTATGTA GGCGGTGCTA CAGAGTTCTT GAAGTGGTGG CCTAACTACG
      CTCCATACAT CCGCCACGAT GTCTCAAGAA CTTCACCACC GGATTGATGC
4951  GCTACACTAG AAGGACAGTA TTTGGTATCT GCGCTCTGCT GAAGCCAGTT
      CGATGTGATC TTCCTGTCAT AAACCATAGA CGCGAGACGA CTTCGGTCAA
5001  ACCTTCGGAA AAAGAGTTGG TAGCTCTTGA TCCGGCAAAC AAACCACCGC
      TGGAAGCCTT TTTCTCAACC ATCGAGAACT AGGCCGTTTG TTTGGTGGCG
```

FIGURE 10-5

5

```
5051   TGGTAGCGGT GGTTTTTTTG TTTGCAAGCA GCAGATTACG CGCAGAAAAA
       ACCATCGCCA CCAAAAAAAC AAACGTTCGT CGTCTAATGC GCGTCTTTTT
5101   AAGGATCTCA AGAAGATCCT TTGATCTTTT CTACGGGGTC TGACGCTCAG
       TTCCTAGAGT TCTTCTAGGA AACTAGAAAA GATGCCCCAG ACTGCGAGTC
5151   TGGAACGAAA ACTCACGTTA AGGGATTTTG GTCATGAGAT TATCAAAAAG
       ACCTTGCTTT TGAGTGCAAT TCCCTAAAAC CAGTACTCTA ATAGTTTTTC
5201   GATCTTCACC TAGATCCTTT TAAATTAAAA ATGAAGTTTT AAATCAATCT
       CTAGAAGTGG ATCTAGGAAA ATTTAATTTT TACTTCAÀAA TTTAGTTAGA
5251   AAAGTATATA TGAGTAAACT TGGTCTGACA GTTACCAATG CTTAATCAGT
       TTTCATATAT ACTCATTTGA ACCAGACTGT CAATGGTTAC GAATTAGTCA
5301   GAGGCACCTA TCTCAGCGAT CTGTCTATTT CGTTCATCCA TAGTTGCCTG
       CTCCGTGGAT AGAGTCGCTA GACAGATAAA GCAAGTAGGT ATCAACGGAC
5351   ACTCCCCGTC GTGTAGATAA CTACGATACG GGAGGGCTTA CCATCTGGCC
       TGAGGGGCAG CACATCTATT GATGCTATGC CCTCCCGAAT GGTAGACCGG
5401   CCAGTGCTGC AATGATACCG CGAGACCCAC GCTCACCGGC TCCAGATTTA
       GGTCACGACG TTACTATGGC GCTCTGGGTG CGAGTGGCCG AGGTCTAAAT
5451   TCAGCAATAA ACCAGCCAGC CGGAAGGGCC GAGCGCAGAA GTGGTCCTGC
       AGTCGTTATT TGGTCGGTCG GCCTTCCCGG CTCGCGTCTT CACCAGGACG
5501   AACTTTATCC GCCTCCATCC AGTCTATTAA TTGTTGCCGG GAAGCTAGAG
       TTGAAATAGG CGGAGGTAGG TCAGATAATT AACAACGGCC CTTCGATCTC
5551   TAAGTAGTTC GCCAGTTAAT AGTTTGCGCA ACGTTGTTGC CATTGCTGCA
       ATTCATCAAG CGGTCAATTA TCAAACGCGT TGCAACAACG GTAACGACGT
5601   GGCATCGTGG TGTCACGCTC GTCGTTTGGT ATGGCTTCAT TCAGCTCCGG
       CCGTAGCACC ACAGTGCGAG CAGCAAACCA TACCGAAGTA AGTCGAGGCC
5651   TTCCCAACGA TCAAGGCGAG TTACATGATC CCCCATGTTG TGCAAAAAAG
       AAGGGTTGCT AGTTCCGCTC AATGTACTAG GGGGTACAAC ACGTTTTTTC
5701   CGGTTAGCTC CTTCGGTCCT CCGATCGTTG TCAGAAGTAA GTTGGCCGCA
       GCCAATCGAG GAAGCCAGGA GGCTAGCAAC AGTCTTCATT CAACCGGCGT
5751   GTGTTATCAC TCATGGTTAT GGCAGCACTG CATAATTCTC TTACTGTCAT
       CACAATAGTG AGTACCAATA CCGTCGTGAC GTATTAAGAG AATGACAGTA
5801   GCCATCCGTA AGATGCTTTT CTGTGACTGG TGAGTACTCA ACCAAGTCAT
       CGGTAGGCAT TCTACGAAAA GACACTGACC ACTCATGAGT TGGTTCAGTA
5851   TCTGAGAATA GTGTATGCGG CGACCGAGTT GCTCTTGCCC GGCGTCAACA
       AGACTCTTAT CACATACGCC GCTGGCTCAA CGAGAACGGG CCGCAGTTGT
5901   CGGGATAATA CCGCGCCACA TAGCAGAACT TTAAAAGTGC TCATCATTGG
       GCCCTATTAT GGCGCGGTGT ATCGTCTTGA AATTTTCACG AGTAGTAACC
5951   AAAACGTTCT TCGGGGCGAA AACTCTCAAG GATCTTACCG CTGTTGAGAT
       TTTTGCAAGA AGCCCCGCTT TTGAGAGTTC CTAGAATGGC GACAACTCTA
6001   CCAGTTCGAT GTAACCCACT CGTGCACCCA ACTGATCTTC AGCATCTTTT
       GGTCAAGCTA CATTGGGTGA GCACGTGGGT TGACTAGAAG TCGTAGAAAA
6051   ACTTTCACCA GCGTTTCTGG GTGAGCAAAA ACAGGAAGGC AAAATGCCGC
       TGAAAGTGGT CGCAAAGACC CACTCGTTTT TGTCCTTCCG TTTTACGGCG
6101   AAAAAAGGGA ATAAGGGCGA CACGGAAATG TTGAATACTC ATACTCTTCC
       TTTTTTCCCT TATTCCCGCT GTGCCTTTAC AACTTATGAG TATGAGAAGG
6151   TTTTTCAATA TTATTGAAGC ATTTATCAGG GTTATTGTCT CATGAGCGGA
       AAAAAGTTAT AATAACTTCG TAAATAGTCC CAATAACAGA GTACTCGCCT
```

FIGURE 10-6

6

```
6201  TACATATTTG  AATGTATTTA  GAAAAATAAA  CAAATAGGGG  TTCCGCGCAC
      ATGTATAAAC  TTACATAAAT  CTTTTTATTT  GTTTATCCCC  AAGGCGCGTG
6251  ATTTCCCCGA  AAAGTGCCAC  CTGACGTCTA  AGAAACCATT  ATTATCATGA
      TAAAGGGGCT  TTTCACGGTG  GACTGCAGAT  TCTTTGGTAA  TAATAGTACT
6301  CATTAACCTA  TAAAAATAGG  CGTATCACGA  GGCCCTTTCG  TCTTCAAGAA
      GTAATTGGAT  ATTTTTATCC  GCATAGTGCT  CCGGGAAAGC  AGAAGTTCTT
6351  TTCTCATGTT  TGACAGCTTA  TCATCGATAA  GCTGATCCTC  ACAGGCCGCA
      AAGAGTACAA  ACTGTCGAAT  AGTAGCTATT  CGACTAGGAG  TGTCCGGCGT
6401  CCCAGCTTTT  CTTCCGTTGC  CCCAGTAGCA  TCTCTGTCTG  GTGACCTTGA
      GGGTCGAAAA  GAAGGCAACG  GGGTCATCGT  AGAGACAGAC  CACTGGAACT
6451  AGAGGAAGAG  GAGGGGTCCC  GAGAATCCCC  ATCCCTACCG  TCCAGCAAAA
      TCTCCTTCTC  CTCCCCAGGG  CTCTTAGGGG  TAGGGATGGC  AGGTCGTTTT
6501  AGGGGGACGA  GGAATTTGAG  GCCTGGCTTG  AGGCTCAGGA  CGCAAATCTT
      TCCCCCTGCT  CCTTAAACTC  CGGACCGAAC  TCCGAGTCCT  GCGTTTAGAA
6551  GAGGATGTTC  AGCGGGAGTT  TTCCGGGCTG  CGAGTAATTG  GTGATGAGGA
      CTCCTACAAG  TCGCCCTCAA  AAGGCCCGAC  GCTCATTAAC  CACTACTCCT
6601  CGAGGATGGT  TCGGAGGATG  GGGAATTTTC  AGACCTGGAT  CTGTCTGACA
      GCTCCTACCA  AGCCTCCTAC  CCCTTAAAAG  TCTGGACCTA  GACAGACTGT
6651  GCGACCATGA  AGGGGATGAG  GGTGGGGGGG  CTGTTGGAGG  GGGCAGGAGT
      CGCTGGTACT  TCCCCTACTC  CCACCCCCCC  GACAACCTCC  CCCGTCCTCA
6701  CTGCACTCCC  TGTATTCACT  GAGCGTCGTC  TAATAAAGAT  GTCTATTGAT
      GACGTGAGGG  ACATAAGTGA  CTCGCAGCAG  ATTATTTCTA  CAGATAACTA
6751  CTCTTTTAGT  GTGAATCATG  TCTGACGAGG  GGCCAGGTAC  AGGACCTGGA
      GAGAAAATCA  CACTTAGTAC  AGACTGCTCC  CCGGTCCATG  TCCTGGACCT
6801  AATGGCCTAG  GAGAGAAGGG  AGACACATCT  GGACCAGAAG  GCTCCGGCGG
      TTACCGGATC  CTCTCTTCCC  TCTGTGTAGA  CCTGGTCTTC  CGAGGCCGCC
6851  CAGTGGACCT  CAAAGAAGAG  GGGGTGATAA  CCATGACGA  GGACGGGGAA
      GTCACCTGGA  GTTTCTTCTC  CCCCACTATT  GGTACCTGCT  CCTGCCCCTT
6901  GAGGACGAGG  ACGAGGAGGC  GGAAGACCAG  GAGCCCCGGG  CGGCTCAGGA
      CTCCTGCTCC  TGCTCCTCCG  CCTTCTGGTC  CTCGGGGCCC  GCCGAGTCCT
6951  TCAGGGCCAA  GACATAGAGA  TGGTGTCCGG  AGACCCCAAA  AACGTCCAAG
      AGTCCCGGTT  CTGTATCTCT  ACCACAGGCC  TCTGGGGTTT  TTGCAGGTTC
7001  TTGCATTGGC  TGCAAAGGGA  CCCACGGTGG  AACAGGAGCA  GGAGCAGGAG
      AACGTAACCG  ACGTTTCCCT  GGGTGCCACC  TTGTCCTCGT  CCTCGTCCTC
7051  CGGGAGGGGC  AGGAGCAGGA  GGGGCAGGAG  CAGGAGGAGG  GGCAGGAGCA
      GCCCTCCCCG  TCCTCGTCCT  CCCCGTCCTC  GTCCTCCTCC  CCGTCCTCGT
7101  GGAGGAGGGG  CAGGAGGGGC  AGGAGGGGCA  GGAGGGGCAG  GAGCAGGAGG
      CCTCCTCCCC  GTCCTCCCCG  TCCTCCCCGT  CCTCCCCGTC  CTCGTCCTCC
7151  AGGGGCAGGA  GCAGGAGGAG  GGGCAGGAGG  GGCAGGAGGG  GCAGGAGCAG
      TCCCCGTCCT  CGTCCTCCTC  CCCGTCCTCC  CCGTCCTCCC  CGTCCTCGTC
7201  GAGGAGGGGC  AGGAGCAGGA  GGAGGGGCAG  GAGGGGCAGG  AGCAGGAGGA
      CTCCTCCCCG  TCCTCGTCCT  CCTCCCCGTC  CTCCCCGTCC  TCGTCCTCCT
7251  GGGGCAGGAG  GGGCAGGAGG  GGCAGGAGCA  GGAGGAGGGG  CAGGAGCAGG
      CCCCGTCCTC  CCCGTCCTCC  CCGTCCTCGT  CCTCCTCCCC  GTCCTCGTCC
7301  AGGAGGGGCA  GGAGGGGCAG  GAGCAGGAGG  AGGGGCAGGA  GGGGCAGGAG
      TCCTCCCCGT  CCTCCCCGTC  CTCGTCCTCC  TCCCCGTCCT  CCCCGTCCTC
```

FIGURE 10-7

7

```
7351  GGGCAGGAGC AGGAGGAGGG GCAGGAGCAG GAGGGGCAGG AGGGGCAGGA
      CCCGTCCTCG TCCTCCTCCC CGTCCTCGTC CTCCCCGTCC TCCCCGTCCT
7401  GGGGCAGGAG CAGGAGGGGC AGGAGCAGGA GGAGGGGCAG GAGGGGCAGG
      CCCCGTCCTC GTCCTCCCCG TCCTCGTCCT CCTCCCCGTC CTCCCCGTCC
7451  AGGGGCAGGA GCAGGAGGGG CAGGAGCAGG AGGGGCAGGA GCAGGAGGGG
      TCCCCGTCCT CGTCCTCCCC GTCCTCGTCC TCCCCGTCCT CGTCCTCCCC
7501  CAGGAGCAGG AGGGGCAGGA GGGGCAGGAG CAGGAGGGGC AGGAGGGGCA
      GTCCTCGTCC TCCCCGTCCT CCCCGTCCTC GTCCTCCCCG TCCTCCCCGT
7551  GGAGCAGGAG GGGCAGGAGG GGCAGGAGCA GGAGGAGGGG CAGGAGGGGC
      CCTCGTCCTC CCCGTCCTCC CCGTCCTCGT CCTCCTCCCC GTCCTCCCCG
7601  AGGAGCAGGA GGAGGGGCAG GAGGGGCAGG AGCAGGAGGG GCAGGAGGGG
      TCCTCGTCCT CCTCCCCGTC CTCCCCGTCC TCGTCCTCCC CGTCCTCCCC
7651  CAGGAGCAGG AGGGGCAGGA GGGGCAGGAG CAGGAGGGGC AGGAGGGGCA
      GTCCTCGTCC TCCCCGTCCT CCCCGTCCTC GTCCTCCCCG TCCTCCCCGT
7701  GGAGCAGGAG GAGGGGCAGG AGCAGGAGGG GCAGGAGCAG GAGGTGGAGG
      CCTCGTCCTC CTCCCCGTCC TCGTCCTCCC CGTCCTCGTC CTCCACCTCC
7751  CCGGGGTCGA GGAGGCAGTG GAGGCCGGGG TCGAGGAGGT AGTGGAGGCC
      GGCCCCAGCT CCTCCGTCAC CTCCGGCCCC AGCTCCTCCA TCACCTCCGG
7801  GGGGTCGAGG AGGTAGTGGA GGCCGCCGGG GTAGAGGACG TGAAAGAGCC
      CCCCAGCTCC TCCATCACCT CCGGCGGCCC CATCTCCTGC ACTTTCTCGG
7851  AGGGGGGGAA GTCGTGAAAG AGCCAGGGGG AGAGGTCGTG GACGTGGAGA
      TCCCCCCCTT CAGCACTTTC TCGGTCCCCC TCTCCAGCAC CTGCACCTCT
7901  AAAGAGGCCC AGGAGTCCCA GTAGTCAGTC ATCATCATCC GGGTCTCCAC
      TTTCTCCGGG TCCTCAGGGT CATCAGTCAG TAGTAGTAGG CCCAGAGGTG
7951  CGCGCAGGCC CCCTCCAGGT AGAAGGCCAT TTTTCCACCC TGTAGGGGAA
      GCGCGTCCGG GGGAGGTCCA TCTTCCGGTA AAAAGGTGGG ACATCCCCTT
8001  GCCGATTATT TTGAATACCA CCAAGAAGGT GGCCCAGATG GTGAGCCTGA
      CGGCTAATAA AACTTATGGT GGTTCTTCCA CCGGGTCTAC CACTCGGACT
8051  CGTGCCCCCG GGAGCGATAG AGCAGGGCCC CGCAGATGAC CCAGGAGAAG
      GCACGGGGGC CCTCGCTATC TCGTCCCGGG GCGTCTACTG GGTCCTCTTC
8101  GCCCAAGCAC TGGACCCCGG GGTCAGGGTG ATGGAGGCAG GCGCAAAAAA
      CGGGTTCGTG ACCTGGGGCC CCAGTCCCAC TACCTCCGTC CGCGTTTTTT
8151  GGAGGGTGGT TTGGAAAGCA TCGTGGTCAA GGAGGTTCCA ACCCGAAATT
      CCTCCCACCA AACCTTTCGT AGCACCAGTT CCTCCAAGGT TGGGCTTTAA
8201  TGAGAACATT GCAGAAGGTT TAAGAGCTCT CCTGGCTAGG AGTCACGTAG
      ACTCTTGTAA CGTCTTCCAA ATTCTCGAGA GGACCGATCC TCAGTGCATC
8251  AAAGGACTAC CGACGAAGGA ACTTGGGTCG CCGGTGTGTT CGTATATGGA
      TTTCCTGATG GCTGCTTCCT TGAACCCAGC GGCCACACAA GCATATACCT
8301  GGTAGTAAGA CCTCCCTTTA CAACCTAAGG CGAGGAACTG CCCTTGCTAT
      CCATCATTCT GGAGGGAAAT GTTGGATTCC GCTCCTTGAC GGGAACGATA
8351  TCCACAATGT CGTCTTACAC CATTGAGTCG TCTCCCCTTT GGAATGGCCC
      AGGTGTTACA GCAGAATGTG GTAACTCAGC AGAGGGGAAA CCTTACCGGG
8401  CTGGACCCGG CCCACAACCT GGCCCGCTAA GGGAGTCCAT TGTCTGTTAT
      GACCTGGGCC GGGTGTTGGA CCGGGCGATT CCCTCAGGTA ACAGACAATA
8451  TTCATGGTCT TTTTACAAAC TCATATATTT GCTGAGGTTT TGAAGGATGC
      AAGTACCAGA AAAATGTTTG AGTATATAAA CGACTCCAAA ACTTCCTACG
```

FIGURE 10-8

```
8501  GATTAAGGAC CTTGTTATGA CAAAGCCCGC TCCTACCTGC AATATCAGGG
      CTAATTCCTG GAACAATACT GTTTCGGGCG AGGATGGACG TTATAGTCCC
8551  TGACTGTGTG CAGCTTTGAC GATGGAGTAG ATTTGCCTCC CTGGTTTCCA
      ACTGACACAC GTCGAAACTG CTACCTCATC TAAACGGAGG GACCAAAGGT
8601  CCTATGGTGG AAGGGGCTGC CGCGGAGGGT GATGACGGAG ATGACGGAGA
      GGATACCACC TTCCCCGACG GCGCCTCCCA CTACTGCCTC TACTGCCTCT
8651  TGAAGGAGGT GATGGAGATG AGGGTGAGGA AGGGCAGGAG TGATGTAACT
      ACTTCCTCCA CTACCTCTAC TCCCACTCCT TCCCGTCCTC ACTACATTGA
8701  TGTTAGGAGA CGCCCTCAAT CGTATTAAAA GCCGTGTATT CCCCCGCACT
      ACAATCCTCT GCGGGAGTTA GCATAATTTT CGGCACATAA GGGGGCGTGA
8751  AAAGAATAAA TCCCCAGTAG ACATCATGCG TGCTGTTGGT GTATTTCTGG
      TTTCTTATTT AGGGGTCATC TGTAGTACGC ACGACAACCA CATAAAGACC
8801  CCATCTGTCT TGTCACCATT TTCGTCCTCC CAACATGGGG CAATTGGGCA
      GGTAGACAGA ACAGTGGTAA AAGCAGGAGG GTTGTACCCC GTTAACCCGT
8851  TACCCATGTT GTCACGTCAC TCAGCTCCGC GCTCAACACC TTCTCGCGTT
      ATGGGTACAA CAGTGCAGTG AGTCGAGGCG CGAGTTGTGG AAGAGCGCAA
8901  GGAAAACATT AGCGACATTT ACCTGGTGAG CAATCAGACA TGCGACGGCT
      CCTTTTGTAA TCGCTGTAAA TGGACCACTC GTTAGTCTGT ACGCTGCCGA
8951  TTAGCCTGGC CTCCTTAAAT TCACCTAAGA ATGGGAGCAA CCAGCATGCA
      AATCGGACCG GAGGAATTTA AGTGGATTCT TACCCTCGTT GGTCGTACGT
9001  GGAAAAGGAC AAGCAGCGAA AATTCACGCC CCCTTGGGAG GTGGCGGCAT
      CCTTTTCCTG TTCGTCGCTT TTAAGTGCGG GGGAACCCTC CACCGCCGTA
9051  ATGCAAAGGA TAGCACTCCC ACTCTACTAC TGGGTATCAT ATGCTGACTG
      TACGTTTCCT ATCGTGAGGG TGAGATGATG ACCCATAGTA TACGACTGAC
9101  TATATGCATG AGGATAGCAT ATGCTACCCG GATACAGATT AGGATAGCAT
      ATATACGTAC TCCTATCGTA TACGATGGGC CTATGTCTAA TCCTATCGTA
9151  ATACTACCCA GATATAGATT AGGATAGCAT ATGCTACCCA GATATAGATT
      TATGATGGGT CTATATCTAA TCCTATCGTA TACGATGGGT CTATATCTAA
9201  AGGATAGCCT ATGCTACCCA GATATAAATT AGGATAGCAT ATACTACCCA
      TCCTATCGGA TACGATGGGT CTATATTTAA TCCTATCGTA TATGATGGGT
9251  GATATAGATT AGGATAGCAT ATGCTACCCA GATATAGATT AGGATAGCCT
      CTATATCTAA TCCTATCGTA TACGATGGGT CTATATCTAA TCCTATCGGA
9301  ATGCTACCCA GATATAGATT AGGATAGCAT ATGCTACCCA GATATAGATT
      TACGATGGGT CTATATCTAA TCCTATCGTA TACGATGGGT CTATATCTAA
9351  AGGATAGCAT ATGCTATCCA GATATTTGGG TAGTATATGC TACCCAGATA
      TCCTATCGTA TACGATAGGT CTATAAACCC ATCATATACG ATGGGTCTAT
9401  TAAATTAGGA TAGCATATAC TACCCTAATC TCTATTAGGA TAGCATATGC
      ATTTAATCCT ATCGTATATG ATGGGATTAG AGATAATCCT ATCGTATACG
9451  TACCCGGATA CAGATTAGGA TAGCATATAC TACCCAGATA TAGATTAGGA
      ATGGGCCTAT GTCTAATCCT ATCGTATATG ATGGGTCTAT ATCTAATCCT
9501  TAGCATATGC TACCCAGATA TAGATTAGGA TAGCCTATGC TACCCAGATA
      ATCGTATACG ATGGGTCTAT ATCTAATCCT ATCGGATACG ATGGGTCTAT
9551  TAAATTAGGA TAGCATATAC TACCCAGATA TAGATTAGGA TAGCATATGC
      ATTTAATCCT ATCGTATATG ATGGGTCTAT ATCTAATCCT ATCGTATACG
9601  TACCCAGATA TAGATTAGGA TAGCCTATGC TACCCAGATA TAGATTAGGA
      ATGGGTCTAT ATCTAATCCT ATCGGATACG ATGGGTCTAT ATCTAATCCT
```

FIGURE 10-9

```
 9651   TAGCATATGC TATCCAGATA TTTGGGTAGT ATATGCTACC CATGGCAACA
        ATCGTATACG ATAGGTCTAT AAACCCATCA TATACGATGG GTACCGTTGT
 9701   TTAGCCCACC GTGCTCTCAG CGACCTCGTG AATATGAGGA CCAACAACCC
        AATCGGGTGG CACGAGAGTC GCTGGAGCAC TTATACTCCT GGTTGTTGGG
 9751   TGTGCTTGGC GCTCAGGCGC AAGTGTGTGT AATTTGTCCT CCAGATCGCA
        ACACGAACCG CGAGTCCGCG TTCACACACA TTAAACAGGA GGTCTAGCGT
 9801   GCAATCGCGC CCCTATCTTG GCCCGCCCAC CTACTTATGC AGGTATTCCC
        CGTTAGCGCG GGGATAGAAC CGGGCGGGTG GATGAATACG TCCATAAGGG
 9851   CGGGGTGCCA TTAGTGGTTT TGTGGGCAAG TGGTTTGACC GCAGTGGTTA
        GCCCCACGGT AATCACCAAA ACACCCGTTC ACCAAACTGG CGTCACCAAT
 9901   GCGGGGTTAC AATCAGCCAA GTTATTACAC CCTTATTTTA CAGTCCAAAA
        CGCCCCAATG TTAGTCGGTT CAATAATGTG GGAATAAAAT GTCAGGTTTT
 9951   CCGCAGGGCG GCGTGTGGGG GCTGACGCGT GCCCCCACTC CACAATTTCA
        GGCGTCCCGC CGCACACCCC CGACTGCGCA CGGGGGTGAG GTGTTAAAGT
10001   AAAAAAAGAG TGGCCACTTG TCTTTGTTTA TGGGCCCCAT TGGCGTGGAG
        TTTTTTTCTC ACCGGTGAAC AGAAACAAAT ACCCGGGGTA ACCGCACCTC
10051   CCCCGTTTAA TTTTCGGGGG TGTTAGAGAC AACCAGTGGA GTCCGCTGCT
        GGGGCAAATT AAAAGCCCCC ACAATCTCTG TTGGTCACCT CAGGCGACGA
10101   GTCGGCGTCC ACTCTCTTTC CCCTTGTTAC AAATAGAGTG TAACAACATG
        CAGCCGCAGG TGAGAGAAAG GGGAACAATG TTTATCTCAC ATTGTTGTAC
10151   GTTCACCTGT CTTGGTCCCT GCCTGGGACA CATCTTAATA ACCCCAGTAT
        CAAGTGGACA GAACCAGGGA CGGACCCTGT GTAGAATTAT TGGGGTCATA
10201   CATATTGCAC TAGGATTATG TGTTGCCCAT AGCCATAAAT TCGTGTGAGA
        GTATAACGTG ATCCTAATAC ACAACGGGTA TCGGTATTTA AGCACACTCT
10251   TGGACATCCA GTCTTTACGG CTTGTCCCCA CCCCATGGAT TTCTATTGTT
        ACCTGTAGGT CAGAAATGCC GAACAGGGGT GGGGTACCTA AAGATAACAA
10301   AAAGATATTC AGAATGTTTC ATTCCTACAC TAGTATTTAT TGCCCAAGGG
        TTTCTATAAG TCTTACAAAG TAAGGATGTG ATCATAAATA ACGGGTTCCC
10351   GTTTGTGAGG GTTATATTGG TGTCATAGCA CAATGCCACC ACTGAACCCC
        CAAACACTCC CAATATAACC ACAGTATCGT GTTACGGTGG TGACTTGGGG
10401   CCGTCCAAAT TTTATTCTGG GGGCGTCACC TGAAACCTTG TTTTCGAGCA
        GGCAGGTTTA AAATAAGACC CCCGCAGTGG ACTTTGGAAC AAAAGCTCGT
10451   CCTCACATAC ACCTTACTGT TCACAACTCA GCAGTTATTC TATTAGCTAA
        GGAGTGTATG TGGAATGACA AGTGTTGAGT CGTCAATAAG ATAATCGATT
10501   ACGAAGGAGA ATGAAGAAGC AGGCGAAGAT TCAGGAGAGT TCACTGCCCG
        TGCTTCCTCT TACTTCTTCG TCCGCTTCTA AGTCCTCTCA AGTGACGGGC
10551   CTCCTTGATC TTCAGCCACT GCCCTTGTGA CTAAAATGGT TCACTACCCT
        GAGGAACTAG AAGTCGGTGA CGGGAACACT GATTTTACCA AGTGATGGGA
10601   CGTGGAATCC TGACCCCATG TAAATAAAAC CGTGACAGCT CATGGGGTGG
        GCACCTTAGG ACTGGGGTAC ATTTATTTTG GCACTGTCGA GTACCCCACC
10651   GAGATATCGC TGTTCCTTAG GACCCTTTTA CTAACCCTAA TTCGATAGCA
        CTCTATAGCG ACAAGGAATC CTGGGAAAAT GATTGGGATT AAGCTATCGT
10701   TATGCTTCCC GTTGGGTAAC ATATGCTATT GAATTAGGGT TAGTCTGGAT
        ATACGAAGGG CAACCCATTG TATACGATAA CTTAATCCCA ATCAGACCTA
10751   AGTATATACT ACTACCCGGG AAGCATATGC TACCCGTTTA GGGTTAACAA
        TCATATATGA TGATGGGCCC TTCGTATACG ATGGGCAAAT CCCAATTGTT
```

FIGURE 10-10

10

```
10801   GGGGGCCTTA TAAACACTAT TGCTAATGCC CTCTTGAGGG TCCGCTTATC
        CCCCCGGAAT ATTTGTGATA ACGATTACGG GAGAACTCCC AGGCGAATAG
10851   GGTAGCTACA CAGGCCCCTC TGATTGACGT TGGTGTAGCC TCCCGTAGTC
        CCATCGATGT GTCCGGGGAG ACTAACTGCA ACCACATCGG AGGGCATCAG
10901   TTCCTGGGCC CCTGGGAGGT ACATGTCCCC CAGCATTGGT GTAAGAGCTT
        AAGGACCCGG GGACCCTCCA TGTACAGGGG GTCGTAACCA CATTCTCGAA
10951   CAGCCAAGAG TTACACATAA AGGCAATGTT GTGTTGCAGT CCACAGACTG
        GTCGGTTCTC AATGTGTATT TCCGTTACAA CACAACGTCA GGTGTCTGAC
11001   CAAAGTCTGC TCCAGGATGA AAGCCACTCA GTGTTGGCAA ATGTGCACAT
        GTTTCAGACG AGGTCCTACT TTCGGTGAGT CACAACCGTT TACACGTGTA
11051   CCATTTATAA GGATGTCAAC TACAGTCAGA GAACCCCTTT GTGTTTGGTC
        GGTAAATATT CCTACAGTTG ATGTCAGTCT CTTGGGGAAA CACAAACCAG
11101   CCCCCCCGTG TCACATGTGG AACAGGGCCC AGTTGGCAAG TTGTACCAAC
        GGGGGGGCAC AGTGTACACC TTGTCCCGGG TCAACCGTTC AACATGGTTG
11151   CAACTGAAGG GATTACATGC ACTGCCCCGC GAAGAAGGGG CAGAGATGCC
        GTTGACTTCC CTAATGTACG TGACGGGGCG CTTCTTCCCC GTCTCTACGG
11201   GTAGTCAGGT TTAGTTCGTC CGGCGGCGGG GCTCTAGAGT CGACCGGTCA
        CATCAGTCCA AATCAAGCAG GCCGCCGCCC CGAGATCTCA GCTGGCCAGT
11251   TGGCTGCGCC CCGACACCCG CCAACACCCG CTGACGCGCC CTGACGGGCT
        ACCGACGCGG GGCTGTGGGC GGTTGTGGGC GACTGCGCGG GACTGCCCGA
11301   TGTCTGCTCC CGGCATCCGC TTACAGACAA GCTGTGACCG TCTCCGGGAG
        ACAGACGAGG GCCGTAGGCG AATGTCTGTT CGACACTGGC AGAGGCCCTC
11351   CTGCATGTGT CAGAGGTTTT CACCGTCATC ACCGAAACGC GCGAGGCAGC
        GACGTACACA GTCTCCAAAA ,GTGGCAGTAG TGGCTTTGCG CGCTCCGTCG
11401   CGGATCATAA TCAGCCATAC CACATTTGTA GAGGTTTTAC TTGCTTTAAA
        GCCTAGTATT AGTCGGTATG GTGTAAACAT CTCCAAAATG AACGAAATTT
11451   AAACCTCCCC ACCTCCCCCT GAACCTGAAA CATAAAATGA ATGCAATTGT
        TTTGGAGGGG TGGAGGGGGA CTTGGACTTT GTATTTTACT TACGTTAACA
11501   TGTTGTTAAC TTGTTTATTG CAGCTTATAA TGGTTACAAA TAAAGCAATA
        ACAACAATTG AACAAATAAC GTCGAATATT ACCAATGTTT ATTTCGTTAT
11551   GCATCACAAA TTTCACAAAT AAAGCATTTT TTTCACTGCA TTCTAGTTGT
        CGTAGTGTTT AAAGTGTTTA TTTCGTAAAA AAAGTGACGT AAGATCAACA
11601   GGTTTGTCCA AACTCATCAA TGTATCTTAT CATGTCTGGA TCCCACGTGC
        CCAAACAGGT TTGAGTAGTT ACATAGAATA GTACAGACCT AGGGTGCACG
11651   AGGCGGGGAG GCGGCCCAAA GGGAGATCCG ACTCGTCTGA GGGCGAAGGC
        TCCGCCCCTC CGCCGGGTTT CCCTCTAGGC TGAGCAGACT CCCGCTTCCG
11701   GAAGACGCGG AAGAGGCCGC AGAGCCGGCA GCAGGCCGCG GGAAGGAAGG
        CTTCTGCGCC TTCTCCGGCG TCTCGGCCGT CGTCCGGCGC CCTTCCTTCC
11751   TCCGCTGGAT TGAGGGCCGA AGGGACGTAG CAGAAGGACG TCCCGCGCAG
        AGGCGACCTA ACTCCCGGCT TCCCTGCATC GTCTTCCTGC AGGGCGCGTC
11801   AATCCAGGTG GCAACACAGG CGAGCAGCCA AGGAAAGGAC GATGATTTCC
        TTAGGTCCAC CGTTGTGTCC GCTCGTCGGT TCCTTTCCTG CTACTAAAGG
11851   CCGACAACAC CACGGAATTG TCAGTGCCCA ACAGCCGAGC CCCTGTCCAG
        GGCTGTTGTG GTGCCTTAAC AGTCACGGGT TGTCGGCTCG GGGACAGGTC
11901   CAGCGGGCAA GGCAGGCGGC GATGAGTTCC GCCGTGGCAA TAGGGAGGGG
        GTCGCCCGTT CCGTCCGCCG CTACTCAAGG CGGCACCGTT ATCCCTCCCC
```

FIGURE 10-11

```
11951   GAAAGCGAAA GTCCCGGAAA GGAGCTGACA GGTGGTGGCA ATGCCCCAAC
        CTTTCGCTTT CAGGGCCTTT CCTCGACTGT CCACCACCGT TACGGGGTTG
12001   CAGTGGGGGT TGCGTCAGCA AACACAGTGC ACACCACGCC ACGTTGCCTG
        GTCACCCCCA ACGCAGTCGT TTGTGTCACG TGTGGTGCGG TGCAACGGAC
12051   ACAACGGGCC ACAACTCCTC ATAAAGAGAC AGCAACCAGG ATTTATACAA
        TGTTGCCCGG TGTTGAGGAG TATTTCTCTG TCGTTGGTCC TAAATATGTT
12101   GGAGGAGAAA ATGAAAGCCA TACGGGAAGC AATAGCATGA TACAAAGGCA
        CCTCCTCTTT TACTTTCGGT ATGCCCTTCG TTATCGTACT ATGTTTCCGT
12151   TTAAAGCAGC GTATCCACAT AGCGTAAAAG GAGCAACATA GTTAAGAATA
        AATTTCGTCG CATAGGTGTA TCGCATTTTC CTCGTTGTAT CAATTCTTAT
12201   CCAGTCAATC TTTCACAAAT TTTGTAATCC AGAGGTTGAT TC
        GGTCAGTTAG AAAGTGTTTA AAACATTAGG TCTCCAACTA AG
```

FIGURE 10-12

12

FIG 11 (1-17)

pCEP4W/hep36

```
  1   TCGAGCGGCC GCTTTAAACT CAATGGTGAT GGTGATGATG ACCGGTACGC
      AGCTCGCCGG CGAAATTTGA GTTACCACTA CCACTACTAC TGGCCATGCG
 -3                        H  H  H   H  H  H  G   T  R  T

 51   GTAGAATCGA GACCGAGGAG AGGGTTAGGG ATAGGCTTAC CGAATTCGAG
      CATCTTAGCT CTGGCTCCTC TCCCAATCCC TATCCGAATG GCTTAAGCTC
 -3    ·T  S  D  L  G  L   L  P  N  P  I  P  K  G   F  E  L

101   CTGGGTCACC ATGCCGCTGG CTTCGGAGTG AGTCTTTATG GCCTGGAAGA
      GACCCAGTGG TACGGCGACC GAAGCCTCAC TCAGAAATAC CGGACCTTCT
 -3    Q  T  V  M  G  S   A  E  S  H  T  K  I  A   Q  F  I·

151   TCCACTCCCG GAAGTCACTG ACTTTGGTGT AGACGCCTGG CTTCTGGGCC
      AGGTGAGGGC CTTCAGTGAC TGAAACCACA TCTGCGGACC GAAGACCCGG
 -3    ·I  W  E  R  F  D   S  V  K  T  Y  V  G  P   K  Q  A  L

201   AGGGCACAGC CAGTGCCCCA ACTCACAATG CCACACAGCC GCCAACGTGG
      TCCCGTGTCG GTCACGGGGT TGAGTGTTAC GGTGTGTCGG CGGTTGCACC
 -3    ·L  A  C  G  T  G   W  S  V  I  G  C  L  R   W  R  P

251   CGTCCGAGAG ATGCTGTCCT CACACACAAA GGGACCACCG CTGTCGCCCT
      GCAGGCTCTC TACGACAGGA GTGTGTGTTT CCCTGGTGGC GACAGCGGGA
 -3    T  R  S  I   S  D  E  C  V  F  P  G  G  S   D  G  Q·

301   GGCAGGCATC AATGCCACCC TCGGGGTAGC CAGCACAGAA CATCTTGGGC
      CCGTCCGTAG TTACGGTGGG AGCCCCATCG GTCGTGTCTT GTAGAACCCG
 -3    ·Q  C  A  D  I  G   G  E  P  Y  G  A  C  F   M  K  P  K

351   TTGATCTGGT TTCCATAGAA ATCAGCGCCA TTGCAGACAT CATTGCTGAT
      AACTAGACCA AAGGTATCTT TAGTCGCGGT AACGTCTGTA GTAACGACTA
 -3    ·K  I  Q  N  G  Y   F  D  A  G  N  C  V  D   N  S  I

401   TATGGGGACT CGAGCCTCCT GGAGTACCCC GGCCTGTTGG CCATAGTACT
      ATACCCCTGA GCTCGGAGGA CCTCATGGGG CCGGACAACC GGTATCATGA
 -3    I  P  V  R   A  E  Q  L  V  G  A  Q  Q  G   Y  Y  Q·

451   GCGTGTTGCC CCAGCCCGTC ACGGTACAGA TCTTGCCATC CACCAGGGCC
      CGCACAACGG GGTCGGGCAG TGCCATGTCT AGAACGGTAG GTGGTCCCGG
 -3    ·Q  T  N  G  W  G   T  V  T  C  I  K  G  D   V  L  A  Q

501   TGGCCGGCAG CTGGGAGGCA CACAGGCTGG ATGTATTCTG TGAGGGGCAG
      ACCGGCCGTC GACCCTCCGT GTGTCCGACC TACATAAGAC ACTCCCCGTC
 -3    ·Q  G  A  A  P  L   C  V  P  Q  I  Y  E  T   L  P  L

551   GGGACTGGAG AGGTGGACCA GGGCAATATC GTTGCTGTTC TCCTCGCTGT
      CCCTGACCTC TCCACCTGGT CCCGTTATAG CAACGACAAG AGGAGCGACA
 -3    P  S  S  L   H  V  L  A  I  D  N  S  N  E   E  S  N·

601   TGGGGTCCCG AAAGGGAAGA TAGCCCCCGT GGTAGACCAC AGCCTGCACC
      ACCCCAGGGC TTTCCCTTCT ATCGGGGGCA CCATCTGGTG TCGGACGTGG
 -3    ·N  P  D  R  F  P   L  Y  G  G  H  Y  V  V   A  Q  V  G

651   CCCAGCTGCA GACCGTGGGG AGAGGCCTGG GCCACGGCAC CGGCAAACAC
      GGGTCGACGT CTGGCACCCC TCTCCGGACC CGGTGCCGTG GCCGTTTGTG
 -3    ·G  L  Q  L  G  H   P  S  A  Q  A  V  A  G   A  F  V

701   TCGCCATCGG GACAGGACCC GGTTCCGCTC CGGGAAGCAG TGGGCGGCTG
      AGCGGTAGCC CTGTCCTGGG CCAAGGCGAG GCCCTTCGTC ACCCGCCGAC
 -3    R  W  R  S   L  V  R  N  R  E  P  F  C  H   A  A  T·
```

FIGURE 11-1

Fig 11

```
 751  TCAGCACCCA GTCCCCGGAG AGCAGGGATC CCCCACAGAG GTGTGCTCCA
       AGTCGTGGGT CAGGGGCCTC TCGTCCCTAG GGGGTGTCTC CACACGAGGT
  -3  ·T  L   V   W   D   G   S   L   L   S   G   G   C   L   H   A   G   D
 801  TCATAGCGAA GGCTGACTTG CCACGGCCAC CGGCCCAAGC TGGTGTCCCG
       AGTATCGCTT CCGACTGAAC GGTGCCGGTG GCCGGGTTCG ACCACAGGGC
  -3  ·D  Y   R   L   S   V   Q   W   P   W   R   G   L   S   T   D   R
 851  GCCTCCCACG ATCTTGTCGT CGTCGTCCAC GGGCAGCTTC CTGCGGCCAC
       CGGAGGGTGC TAGAACAGCA GCAGCAGGTG CCCGTCGAAG GACGCCGGTG
  -3   G   G   V   I   K   D   D   D   D   V   P   L   K   R   R   G   C·
 901  AGTCTTGGTC CGGAGCGTCA CCAGTGGAAC CTGGAACCCA GAGCAGCAGT
       TCAGAACCAG GCCTCGCAGT GGTCACCTTG GACCTTGGGT CTCGTCGTCA
  -3  ·C  D   Q   D   P   A   D   G   T   S   G   P   V   W   L   L   L   V
 951  ACCCATAGCA GGAGTGTGTC TGTCTCCATG GTGGCGATCT GGTACCCAGC
       TGGGTATCGT CCTCACACAG ACAGAGGTAC CACCGCTAGA CCATGGGTCG
  -3  ·V  W   L   L   L   T   D   T   E   M
1001  TTCTAGAGAT CTGACGGTTC ACTAAACGAG CTCTGCTTAT ATAGACCTCC
       AAGATCTCTA GACTGCCAAG TGATTTGCTC GAGACGAATA TATCTGGAGG
1051  CACCGTACAC GCCTACCGCC CATTTGCGTC AACGGGGCGG GGTTATTACG
       GTGGCATGTG CGGATGGCGG GTAAACGCAG TTGCCCCGCC CCAATAATGC
1101  ACATTTTGGA AAGTCCCGTT GATTTGGTG CCAAAACAAA CTCCCATTGA
       TGTAAAACCT TTCAGGGCAA CTAAAACCAC GGTTTTGTTT GAGGGTAACT
1151  CGTCAATGGG GTGGAGACTT GGAAATCCCC GTGAGTCAAA CCGCTATCCA
       GCAGTTACCC CACCTCTGAA CCTTTAGGGG CACTCAGTTT GGCGATAGGT
1201  CGCCCATTGG TGTACTGCCA AAACCGCATC ACCATGGTAA TAGCGATGAC
       GCGGGTAACC ACATGACGGT TTTGGCGTAG TGGTACCATT ATCGCTACTG
1251  TAATACGTAG ATGTACTGCC AAGTAGGAAA GTCCCGTAAG GTCATGTACT
       ATTATGCATC TACATGACGG TTCATCCTTT CAGGGCATTC CAGTACATGA
1301  GGGCATAATG CCAGGCGGGC CATTTACCGT CATTGACGTC AATAGGGGGC
       CCCGTATTAC GGTCCGCCCG GTAAATGGCA GTAACTGCAG TTATCCCCCG
1351  GGACTTGGCA TATGATACAC TTGATGTACT GCCAAGTGGG CAGTTTACCG
       CCTGAACCGT ATACTATGTG AACTACATGA CGGTTCACCC GTCAAATGGC
1401  TAAATACTCC ACCCATTGAC GTCAATGGAA AGTCCCTATT GGCGTTACTA
       ATTTATGAGG TGGGTAACTG CAGTTACCTT TCAGGGATAA CCGCAATGAT
1451  TGGGAACATA CGTCATTATT GACGTCAATG GCGGGGGGTC GTTGGGCGGT
       ACCCTTGTAT GCAGTAATAA CTGCAGTTAC CGCCCCCAG CAACCCGCCA
1501  CAGCCAGGCG GGCCATTTAC CGTAAGTTAT GTAACGCGGA ACTCCATATA
       GTCGGTCCGC CCGGTAAATG GCATTCAATA CATTGCGCCT TGAGGTATAT
1551  TGGGCTATGA ACTAATGACC CCGTAATTGA TTACTATTAA TAACTAGTCA
       ACCCGATACT TGATTACTGG GGCATTAACT AATGATAATT ATTGATCAGT
1601  ATAATCAATG TCAACATGGC GGTCATATTG ACATGAGCC AATATAAATG
       TATTAGTTAC AGTTGTACCG CCAGTATAAC CTGTACTCGG TTATATTTAC
1651  TACATATTAT GATATAGATA CAACGTATGC AATGGCCAAT AGCCAATATT
       ATGTATAATA CTATATCTAT GTTGCATACG TTACCGGTTA TCGGTTATAA
1701  GATTTATGCT ATATAACCAA TGACTAATAT GGCTAATTGC AATATTGAT
       CTAAATACGA TATATTGGTT ACTGATTATA CCGATTAACG GTTATAACTA
```

FIGURE 11-2

2

```
1751  TCAATGTATA GATCTTCCAT ACCTACCAGT TCTGCGCCTG CAGCAATGCA
      AGTTACATAT CTAGAAGGTA TGGATGGTCA AGACGCGGAC GTCGTTACGT
1801  ACAACGTTGC CCGGATCTGC GATGATAAGC TGTCAAACAT GAGAATTGGT
      TGTTGCAACG GGCCTAGACG CTACTATTCG ACAGTTTGTA CTCTTAACCA
1851  CGACTAGCTT GGCACGCCAG AAATCCGCGC GGTGGTTTTT GGGGGTCGGG
      GCTGATCGAA CCGTGCGGTC TTTAGGCGCG CCACCAAAAA CCCCCAGCCC
1901  GGTGTTTGGC AGCCACAGAC GCCCGGTGTT CGTGTCGCGC CAGTACATGC
      CCACAAACCG TCGGTGTCTG CGGGCCACAA GCACAGCGCG GTCATGTACG
1951  GGTCCATGCC CAGGCCATCC AAAAACCATG GGTCTGTCTG CTCAGTCCAG
      CCAGGTACGG GTCCGGTAGG TTTTTGGTAC CCAGACAGAC GAGTCAGGTC
2001  TCGTGGACCA GACCCCACGC AACGCCCAAA ATAATAACCC CCACGAACCA
      AGCACCTGGT CTGGGGTGCG TTGCGGGTTT TATTATTGGG GGTGCTTGGT
2051  TAAACCATTC CCCATGGGGG ACCCCGTCCC TAACCCACGG GGCCAGTGGC
      ATTTGGTAAG GGGTACCCCC TGGGGCAGGG ATTGGGTGCC CCGGTCACCG
2101  TATGGCAGGG CCTGCCGCCC CGACGTTGGC TGCGAGCCCT GGGCCTTCAC
      ATACCGTCCC GGACGGCGGG GCTGCAACCG ACGCTCGGGA CCCGGAAGTG
2151  CCGAACTTGG GGGGTGGGGT GGGGAAAAGG AAGAAACGCG GGCGTATTGG
      GGCTTGAACC CCCCACCCCA CCCCTTTTCC TTCTTTGCGC CCGCATAACC
2201  CCCCAATGGG GTCTCGGTGG GGTATCGACA GAGTGCCAGC CCTGGGACCG
      GGGGTTACCC CAGAGCCACC CCATAGCTGT CTCACGGTCG GGACCCTGGC
2251  AACCCCGCGT TTATGAACAA ACGACCCAAC ACCCGTGCGT TTTATTCTGT
      TTGGGGCGCA AATACTTGTT TGCTGGGTTG TGGGCACGCA AAATAAGACA
2301  CTTTTTATTG CCGTCATAGC GCGGGTTCCT TCCGGTATTG TCTCCTTCCG
      GAAAAATAAC GGCAGTATCG CGCCCAAGGA AGGCCATAAC AGAGGAAGGC
2351  TGTTTCAGTT AGCCTCCCCC ATCTCCCCTA TTCCTTTGCC CTCGGACGAG
      ACAAAGTCAA TCGGAGGGGG TAGAGGGGAT AAGGAAACGG GAGCCTGCTC
2401  TGCTGGGGCG TCGGTTTCCA CTATCGGCGA GTACTTCTAC ACAGCCATCG
      ACGACCCCGC AGCCAAAGGT GATAGCCGCT CATGAAGATG TGTCGGTAGC
2451  GTCCAGACGG CCGCGCTTCT GCGGGCGATT TGTGTACGCC CGACAGTCCC
      CAGGTCTGCC GGCGCGAAGA CGCCCGCTAA ACACATGCGG GCTGTCAGGG
2501  GGCTCCGGAT CGGACGATTG CGTCGCATCG ACCCTGCGCC CAAGCTGCAT
      CCGAGGCCTA GCCTGCTAAC GCAGCGTAGC TGGGACGCGG GTTCGACGTA
2551  CATCGAAATT GCCGTCAACC AAGCTCTGAT AGAGTTGGTC AAGACCAATG
      GTAGCTTTAA CGGCAGTTGG TTCGAGACTA TCTCAACCAG TTCTGGTTAC
2601  CGGAGCATAT ACGCCCGGAG CCGCGGCGAT CCTGCAAGCT CCGGATGCCT
      GCCTCGTATA TGCGGGCCTC GGCGCCGCTA GGACGTTCGA GGCCTACGGA
2651  CCGCTCGAAG TAGCGCGTCT GCTGCTCCAT ACAAGCCAAC CACGGCCTCC
      GGCGAGCTTC ATCGCGCAGA CGACGAGGTA TGTTCGGTTG GTGCCGGAGG
2701  AGAAGAAGAT GTTGGCGACC TCGTATTGGG AATCCCCGAA CATCGCCTCG
      TCTTCTTCTA CAACCGCTGG AGCATAACCC TTAGGGGCTT GTAGCGGAGC
2751  CTCCAGTCAA TGACCGCTGT TATGCGGCCA TTGTCCGTCA GGACATTGTT
      GAGGTCAGTT ACTGGCGACA ATACGCCGGT AACAGGCAGT CCTGTAACAA
2801  GGAGCCGAAA TCCGCGTGCA CGAGGTGCCG GACTTCGGGG CAGTCCTCGG
      CCTCGGCTTT AGGCGCACGT GCTCCACGGC CTGAAGCCCC GTCAGGAGCC
2851  CCCAAAGCAT CAGCTCATCG AGAGCCTGCG CGACGGACGC ACTGACGGTG
      GGGTTTCGTA GTCGAGTAGC TCTCGGACGC GCTGCCTGCG TGACTGCCAC
```

FIGURE 11-3

3

```
2901  TCGTCCATCA CAGTTTGCCA GTGATACACA TGGGGATCAG CAATCGCGCA
      AGCAGGTAGT GTCAAACGGT CACTATGTGT ACCCCTAGTC GTTAGCGCGT
2951  TATGAAATCA CGCCATGTAG TGTATTGACC GATTCCTTGC GGTCCGAATG
      ATACTTTAGT GCGGTACATC ACATAACTGG CTAAGGAACG CCAGGCTTAC
3001  GGCCGAACCC GCTCGTCTGG CTAAGATCGG CCGCAGCGAT CGCATCCATG
      CCGGCTTGGG CGAGCAGACC GATTCTAGCC GGCGTCGCTA GCGTAGGTAC
3051  GCCTCCGCGA CCGGCTGCAG AACAGCGGGC AGTTCGGTTT CAGGCAGGTC
      CGGAGGCGCT GGCCGACGTC TTGTCGCCCG TCAAGCCAAA GTCCGTCCAG
3101  TTGCAACGTG ACACCCTGTG CACGGCGGGA GATGCAATAG GTCAGGCTCT
      AACGTTGCAC TGTGGGACAC GTGCCGCCCT CTACGTTATC CAGTCCGAGA
3151  CGCTGAATTC CCCAATGTCA AGCACTTCCG GAATCGGGAG CGCGGCCGAT
      GCGACTTAAG GGGTTACAGT TCGTGAAGGC CTTAGCCCTC GCGCCGGCTA
3201  GCAAAGTGCC GATAAACATA ACGATCTTTG TAGAAACCAT CGGCGCAGCT
      CGTTTCACGG CTATTTGTAT TGCTAGAAAC ATCTTTGGTA GCCGCGTCGA
3251  ATTTACCCGC AGGACATATC CACGCCCTCC TACATCGAAG CTGAAAGCAC
      TAAATGGGCG TCCTGTATAG GTGCGGGAGG ATGTAGCTTC GACTTTCGTG
3301  GAGATTCTTC GCCCTCCGAG AGCTGCATCA GGTCGGAGAC GCTGTCGAAC
      CTCTAAGAAG CGGGAGGCTC TCGACGTAGT CCAGCCTCTG CGACAGCTTG
3351  TTTTCGATCA GAAACTTCTC GACAGACGTC GCGGTGAGTT CAGGCTTTTT
      AAAAGCTAGT CTTTGAAGAG CTGTCTGCAG CGCCACTCAA GTCCGAAAAA
3401  CATATCTCAT TGCCCGGGAT CTGCGGCACG CTGTTGACGC TGTTAAGCGG
      GTATAGAGTA ACGGGCCCTA GACGCCGTGC GACAACTGCG ACAATTCGCC
3451  GTCGCTGCAG GGTCGCTCGG TGTTCGAGGC CACACGCGTC ACCTTAATAT
      CAGCGACGTC CCAGCGAGCC ACAAGCTCCG GTGTGCGCAG TGGAATTATA
3501  GCGAAGTGGA CCTGGGACCG CGCCGCCCCG ACTGCATCTG CGTGTTCGAA
      CGCTTCACCT GGACCCTGGC GCGGCGGGGC TGACGTAGAC GCACAAGCTT
3551  TTCGCCAATG ACAAGACGCT GGGCGGGGTT TGTGTCATCA TAGAACTAAA
      AAGCGGTTAC TGTTCTGCGA CCCGCCCCAA ACACAGTAGT ATCTTGATTT
3601  GACATGCAAA TATATTTCTT CCGGGGACAC CGCCAGCAAA CGCGAGCAAC
      CTGTACGTTT ATATAAAGAA GGCCCCTGTG GCGGTCGTTT GCGCTCGTTG
3651  GGGCCACGGG GATGAAGCAG GGCATGGCGG CCGACGCGCT GGGCTACGTC
      CCCGGTGCCC CTACTTCGTC CCGTACCGCC GGCTGCGCGA CCCGATGCAG
3701  TTGCTGGCGT TCGCGACGCG AGGCTGGATG GCCTTCCCCA TTATGATTCT
      AACGACCGCA AGCGCTGCGC TCCGACCTAC CGGAAGGGGT AATACTAAGA
3751  TCTCGCTTCC GGCGGCATCG GGATGCCCGC GTTGCAGGCC ATGCTGTCCA
      AGAGCGAAGG CCGCCGTAGC CCTACGGGCG CAACGTCCGG TACGACAGGT
3801  GGCAGGTAGA TGACGACCAT CAGGGACAGC TTCAAGGATC GCTCGCGGCT
      CCGTCCATCT ACTGCTGGTA GTCCCTGTCG AAGTTCCTAG CGAGCGCCGA
3851  CTTACCAGCC TAACTTCGAT CACTGGACCG CTGATCGTCA CGGCGATTTA
      GAATGGTCGG ATTGAAGCTA GTGACCTGGC GACTAGCAGT GCCGCTAAAT
3901  TGCCGCCTCG GCGAGCACAT GGAACGGGTT GGCATGGATT GTAGGCGCCG
      ACGGCGGAGC CGCTCGTGTA CCTTGCCCAA CCGTACCTAA CATCCGCGGC
3951  CCCTATACCT TGTCTGCCTC CCCGCGTTGC GTCGCGGTGC ATGGAGCCGG
      GGGATATGGA ACAGACGGAG GGGCGCAACG CAGCGCCACG TACCTCGGCC
4001  GCCACCTCGA CCTGAATGGA AGCCGGCGGC ACCTCGCTAA CGGATTCACC
      CGGTGGAGCT GGACTTACCT TCGGCCGCCG TGGAGCGATT GCCTAAGTGG
```

FIGURE 11-4

4

```
4051   ACTCCAAGAA TTGGAGCCAA TCAATTCTTG CGGAGAACTG TGAATGCGCA
       TGAGGTTCTT AACCTCGGTT AGTTAAGAAC GCCTCTTGAC ACTTACGCGT
4101   AACCAACCCT TGGCAGAACA TATCCATCGC GTCCGCCATC TCCAGCAGCC
       TTGGTTGGGA ACCGTCTTGT ATAGGTAGCG CAGGCGGTAG AGGTCGTCGG
4151   GCACGCGGCG CAGCAAAAGG CCAGGAACCG TAAAAAGGCC GCGTTGCTGG
       CGTGCGCCGC GTCGTTTTCC GGTCCTTGGC ATTTTTCCGG CGCAACGACC
4201   CGTTTTTCCA TAGGCTCCGC CCCCCTGACG AGCATCACAA AAATCGACGC
       GCAAAAAGGT ATCCGAGGCG GGGGGACTGC TCGTAGTGTT TTTAGCTGCG
4251   TCAAGTCAGA GGTGGCGAAA CCCGACAGGA CTATAAAGAT ACCAGGCGTT
       AGTTCAGTCT CCACCGCTTT GGGCTGTCCT GATATTTCTA TGGTCCGCAA
4301   TCCCCCTGGA AGCTCCCTCG TGCGCTCTCC TGTTCCGACC CTGCCGCTTA
       AGGGGGACCT TCGAGGGAGC ACGCGAGAGG ACAAGGCTGG GACGGCGAAT
4351   CCGGATACCT GTCCGCCTTT CTCCCTTCGG GAAGCGTGGC GCTTTCTCAT
       GGCCTATGGA CAGGCGGAAA GAGGGAAGCC CTTCGCACCG CGAAAGAGTA
4401   AGCTCACGCT GTAGGTATCT CAGTTCGGTG TAGGTCGTTC GCTCCAAGCT
       TCGAGTGCGA CATCCATAGA GTCAAGCCAC ATCCAGCAAG CGAGGTTCGA
4451   GGGCTGTGTG CACGAACCCC CCGTTCAGCC CGACCGCTGC GCCTTATCCG
       CCCGACACAC GTGCTTGGGG GGCAAGTCGG GCTGGCGACG CGGAATAGGC
4501   GTAACTATCG TCTTGAGTCC AACCCGGTAA GACACGACTT ATCGCCACTG
       CATTGATAGC AGAACTCAGG TTGGGCCATT CTGTGCTGAA TAGCGGTGAC
4551   GCAGCAGCCA CTGGTAACAG GATTAGCAGA GCGAGGTATG TAGGCGGTGC
       CGTCGTCGGT GACCATTGTC CTAATCGTCT CGCTCCATAC ATCCGCCACG
4601   TACAGAGTTC TTGAAGTGGT GGCCTAACTA CGGCTACACT AGAAGGACAG
       ATGTCTCAAG AACTTCACCA CCGGATTGAT GCCGATGTGA TCTTCCTGTC
4651   TATTTGGTAT CTGCGCTCTG CTGAAGCCAG TTACCTTCGG AAAAAGAGTT
       ATAAACCATA GACGCGAGAC GACTTCGGTC AATGGAAGCC TTTTTCTCAA
4701   GGTAGCTCTT GATCCGGCAA ACAAACCACC GCTGGTAGCG GTGGTTTTTT
       CCATCGAGAA CTAGGCCGTT TGTTTGGTGG CGACCATCGC CACCAAAAAA
4751   TGTTTGCAAG CAGCAGATTA CGCGCAGAAA AAAAGGATCT CAAGAAGATC
       ACAAACGTTC GTCGTCTAAT GCGCGTCTTT TTTTCCTAGA GTTCTTCTAG
4801   CTTTGATCTT TTCTACGGGG TCTGACGCTC AGTGGAACGA AAACTCACGT
       GAAACTAGAA AAGATGCCCC AGACTGCGAG TCACCTTGCT TTTGAGTGCA
4851   TAAGGGATTT TGGTCATGAG ATTATCAAAA AGGATCTTCA CCTAGATCCT
       ATTCCCTAAA ACCAGTACTC TAATAGTTTT TCCTAGAAGT GGATCTAGGA
4901   TTTAAATTAA AAATGAAGTT TTAAATCAAT CTAAAGTATA TATGAGTAAA
       AAATTTAATT TTTACTTCAA AATTTAGTTA GATTTCATAT ATACTCATTT
4951   CTTGGTCTGA CAGTTACCAA TGCTTAATCA GTGAGGCACC TATCTCAGCG
       GAACCAGACT GTCAATGGTT ACGAATTAGT CACTCCGTGG ATAGAGTCGC
5001   ATCTGTCTAT TTCGTTCATC CATAGTTGCC TGACTCCCCG TCGTGTAGAT
       TAGACAGATA AAGCAAGTAG GTATCAACGG ACTGAGGGGC AGCACATCTA
5051   AACTACGATA CGGGAGGGCT TACCATCTGG CCCCAGTGCT GCAATGATAC
       TTGATGCTAT GCCCTCCCGA ATGGTAGACC GGGGTCACGA CGTTACTATG
5101   CGCGAGACCC ACGCTCACCG GCTCCAGATT TATCAGCAAT AAACCAGCCA
       GCGCTCTGGG TGCGAGTGGC CGAGGTCTAA ATAGTCGTTA TTTGGTCGGT
5151   GCCGGAAGGG CCGAGCGCAG AAGTGGTCCT GCAACTTTAT CCGCCTCCAT
       CGGCCTTCCC GGCTCGCGTC TTCACCAGGA CGTTGAAATA GGCGGAGGTA
```

FIGURE 11-5

```
5201  CCAGTCTATT AATTGTTGCC GGGAAGCTAG AGTAAGTAGT TCGCCAGTTA
      GGTCAGATAA TTAACAACGG CCCTTCGATC TCATTCATCA AGCGGTCAAT
5251  ATAGTTTGCG CAACGTTGTT GCCATTGCTG CAGGCATCGT GGTGTCACGC
      TATCAAACGC GTTGCAACAA CGGTAACGAC GTCCGTAGCA CCACAGTGCG
5301  TCGTCGTTTG GTATGGCTTC ATTCAGCTCC GGTTCCCAAC GATCAAGGCG
      AGCAGCAAAC CATACCGAAG TAAGTCGAGG CCAAGGGTTG CTAGTTCCGC
5351  AGTTACATGA TCCCCCATGT TGTGCAAAAA AGCGGTTAGC TCCTTCGGTC
      TCAATGTACT AGGGGGTACA ACACGTTTTT TCGCCAATCG AGGAAGCCAG
5401  CTCCGATCGT TGTCAGAAGT AAGTTGGCCG CAGTGTTATC ACTCATGGTT
      GAGGCTAGCA ACAGTCTTCA TTCAACCGGC GTCACAATAG TGAGTACCAA
5451  ATGGCAGCAC TGCATAATTC TCTTACTGTC ATGCCATCCG TAAGATGCTT
      TACCGTCGTG ACGTATTAAG AGAATGACAG TACGGTAGGC ATTCTACGAA
5501  TTCTGTGACT GGTGAGTACT CAACCAAGTC ATTCTGAGAA TAGTGTATGC
      AAGACACTGA CCACTCATGA GTTGGTTCAG TAAGACTCTT ATCACATACG
5551  GGCGACCGAG TTGCTCTTGC CCGGCGTCAA CACGGGATAA TACCGCGCCA
      CCGCTGGCTC AACGAGAACG GGCCGCAGTT GTGCCCTATT ATGGCGCGGT
5601  CATAGCAGAA CTTTAAAAGT GCTCATCATT GGAAAACGTT CTTCGGGGCG
      GTATCGTCTT GAAATTTTCA CGAGTAGTAA CCTTTTGCAA GAAGCCCCGC
5651  AAAACTCTCA AGGATCTTAC CGCTGTTGAG ATCCAGTTCG ATGTAACCCA
      TTTTGAGAGT TCCTAGAATG GCGACAACTC TAGGTCAAGC TACATTGGGT
5701  CTCGTGCACC CAACTGATCT TCAGCATCTT TTACTTTCAC CAGCGTTTCT
      GAGCACGTGG GTTGACTAGA AGTCGTAGAA AATGAAAGTG GTCGCAAAGA
5751  GGGTGAGCAA AAACAGGAAG GCAAAATGCC GCAAAAAAGG GAATAAGGGC
      CCCACTCGTT TTTGTCCTTC CGTTTTACGG CGTTTTTTCC CTTATTCCCG
5801  GACACGGAAA TGTTGAATAC TCATACTCTT CCTTTTTCAA TATTATTGAA
      CTGTGCCTTT ACAACTTATG AGTATGAGAA GGAAAAAGTT ATAATAACTT
5851  GCATTTATCA GGGTTATTGT CTCATGAGCG GATACATATT TGAATGTATT
      CGTAAATAGT CCCAATAACA GAGTACTCGC CTATGTATAA ACTTACATAA
5901  TAGAAAAATA AACAAATAGG GGTTCCGCGC ACATTTCCCC GAAAAGTGCC
      ATCTTTTTAT TTGTTTATCC CCAAGGCGCG TGTAAAGGGG CTTTTCACGG
5951  ACCTGACGTC TAAGAAACCA TTATTATCAT GACATTAACC TATAAAAATA
      TGGACTGCAG ATTCTTTGGT AATAATAGTA CTGTAATTGG ATATTTTTAT
6001  GGCGTATCAC GAGGCCCTTT CGTCTTCAAG AATTCTCATG TTTGACAGCT
      CCGCATAGTG CTCCGGGAAA GCAGAAGTTC TTAAGAGTAC AAACTGTCGA
6051  TATCATCGAT AAGCTGATCC TCACAGGCCG CACCCAGCTT TTCTTCCGTT
      ATAGTAGCTA TTCGACTAGG AGTGTCCGGC GTGGGTCGAA AAGAAGGCAA
6101  GCCCCAGTAG CATCTCTGTC TGGTGACCTT GAAGAGGAAG AGGAGGGGTC
      CGGGGTCATC GTAGAGACAG ACCACTGGAA CTTCTCCTTC TCCTCCCCAG
6151  CCGAGAATCC CCATCCCTAC CGTCCAGCAA AAAGGGGGAC GAGGAATTTG
      GGCTCTTAGG GGTAGGGATG GCAGGTCGTT TTTCCCCCTG CTCCTTAAAC
6201  AGGCCTGGCT TGAGGCTCAG GACGCAAATC TTGAGGATGT TCAGCGGGAG
      TCCGGACCGA ACTCCGAGTC CTGCGTTTAG AACTCCTACA AGTCGCCCTC
6251  TTTTCCGGGC TGCGAGTAAT TGGTGATGAG GACGAGGATG GTTCGGAGGA
      AAAAGGCCCG ACGCTCATTA ACCACTACTC CTGCTCCTAC CAAGCCTCCT
6301  TGGGGAATTT TCAGACCTGG ATCTGTCTGA CAGCGACCAT GAAGGGGATG
      ACCCCTTAAA AGTCTGGACC TAGACAGACT GTCGCTGGTA CTTCCCCTAC
```

FIGURE 11-6

```
6351   AGGGTGGGGG  GGCTGTTGGA  GGGGGCAGGA  GTCTGCACTC  CCTGTATTCA
       TCCCACCCCC  CCGACAACCT  CCCCCGTCCT  CAGACGTGAG  GGACATAAGT
6401   CTGAGCGTCG  TCTAATAAAG  ATGTCTATTG  ATCTCTTTTA  GTGTGAATCA
       GACTCGCAGC  AGATTATTTC  TACAGATAAC  TAGAGAAAAT  CACACTTAGT
6451   TGTCTGACGA  GGGGCCAGGT  ACAGGACCTG  GAAATGGCCT  AGGAGAGAAG
       ACAGACTGCT  CCCCGGTCCA  TGTCCTGGAC  CTTTACCGGA  TCCTCTCTTC
6501   GGAGACACAT  CTGGACCAGA  AGGCTCCGGC  GGCAGTGGAC  CTCAAAGAAG
       CCTCTGTGTA  GACCTGGTCT  TCCGAGGCCG  CCGTCACCTG  GAGTTTCTTC
6551   AGGGGGTGAT  AACCATGGAC  GAGGACGGGG  AAGAGGACGA  GGACGAGGAG
       TCCCCCACTA  TTGGTACCTG  CTCCTGCCCC  TTCTCCTGCT  CCTGCTCCTC
6601   GCGGAAGACC  AGGAGCCCCG  GGCGGCTCAG  GATCAGGGCC  AAGACATAGA
       CGCCTTCTGG  TCCTCGGGGC  CCGCCGAGTC  CTAGTCCCGG  TTCTGTATCT
6651   GATGGTGTCC  GGAGACCCCA  AAAACGTCCA  AGTTGCATTG  GCTGCAAAGG
       CTACCACAGG  CCTCTGGGGT  TTTTGCAGGT  TCAACGTAAC  CGACGTTTCC
6701   GACCCACGGT  GGAACAGGAG  CAGGAGCAGG  AGCGGGAGGG  GCAGGAGCAG
       CTGGGTGCCA  CCTTGTCCTC  GTCCTCGTCC  TCGCCCTCCC  CGTCCTCGTC
6751   GAGGGGCAGG  AGCAGGAGGA  GGGGCAGGAG  CAGGAGGAGG  GGCAGGAGGG
       CTCCCCGTCC  TCGTCCTCCT  CCCCGTCCTC  GTCCTCCTCC  CCGTCCTCCC
6801   GCAGGAGGGG  CAGGAGGGGC  AGGAGCAGGA  GGAGGGGCAG  GAGCAGGAGG
       CGTCCTCCCC  GTCCTCCCCG  TCCTCGTCCT  CCTCCCCGTC  CTCGTCCTCC
6851   AGGGGCAGGA  GGGGCAGGAG  GGGCAGGAGC  AGGAGGAGGG  GCAGGAGCAG
       TCCCCGTCCT  CCCCGTCCTC  CCCGTCCTCG  TCCTCCTCCC  CGTCCTCGTC
6901   GAGGAGGGGC  AGGAGGGGCA  GGAGCAGGAG  GAGGGGCAGG  AGGGGCAGGA
       CTCCTCCCCG  TCCTCCCCGT  CCTCGTCCTC  CTCCCCGTCC  TCCCCGTCCT
6951   GGGGCAGGAG  CAGGAGGAGG  GGCAGGAGCA  GGAGGAGGGG  CAGGAGGGGC
       CCCCGTCCTC  GTCCTCCTCC  CCGTCCTCGT  CCTCCTCCCC  GTCCTCCCCG
7001   AGGAGCAGGA  GGAGGGGCAG  GAGGGGCAGG  AGGGGCAGGA  GCAGGAGGAG
       TCCTCGTCCT  CCTCCCCGTC  CTCCCCGTCC  TCCCCGTCCT  CGTCCTCCTC
7051   GGGCAGGAGC  AGGAGGGGCA  GGAGGGGCAG  GAGGGGCAGG  AGCAGGAGGG
       CCCGTCCTCG  TCCTCCCCGT  CCTCCCCGTC  CTCCCCGTCC  TCGTCCTCCC
7101   GCAGGAGCAG  GAGGAGGGGC  AGGAGGGGCA  GGAGGGGCAG  GAGCAGGAGG
       CGTCCTCGTC  CTCCTCCCCG  TCCTCCCCGT  CCTCCCCGTC  CTCGTCCTCC
7151   GGCAGGAGCA  GGAGGGGCAG  GAGCAGGAGG  GGCAGGAGCA  GGAGGGGCAG
       CCGTCCTCGT  CCTCCCCGTC  CTCGTCCTCC  CCGTCCTCGT  CCTCCCCGTC
7201   GAGGGGCAGG  AGCAGGAGGG  GCAGGAGGGG  CAGGAGCAGG  AGGGGCAGGA
       CTCCCCGTCC  TCGTCCTCCC  CGTCCTCCCC  GTCCTCGTCC  TCCCCGTCCT
7251   GGGGCAGGAG  CAGGAGGAGG  GGCAGGAGGG  GCAGGAGCAG  GAGGAGGGGC
       CCCCGTCCTC  GTCCTCCTCC  CCGTCCTCCC  CGTCCTCGTC  CTCCTCCCCG
7301   AGGAGGGGCA  GGAGCAGGAG  GGGCAGGAGG  GGCAGGAGCA  GGAGGGGCAG
       TCCTCCCCGT  CCTCGTCCTC  CCCGTCCTCC  CCGTCCTCGT  CCTCCCCGTC
7351   GAGGGGCAGG  AGCAGGAGGG  GCAGGAGGGG  CAGGAGCAGG  AGGAGGGGCA
       CTCCCCGTCC  TCGTCCTCCC  CGTCCTCCCC  GTCCTCGTCC  TCCTCCCCGT
7401   GGAGCAGGAG  GGGCAGGAGC  AGGAGGTGGA  GGCCGGGGTC  GAGGAGGCAG
       CCTCGTCCTC  CCCGTCCTCG  TCCTCCACCT  CCGGCCCCAG  CTCCTCCGTC
7451   TGGAGGCCGG  GGTCGAGGAG  GTAGTGGAGG  CCGGGGTCGA  GGAGGTAGTG
       ACCTCCGGCC  CCAGCTCCTC  CATCACCTCC  GGCCCCAGCT  CCTCCATCAC
```

FIGURE 11-7

7

```
7501  GAGGCCGCCG GGGTAGAGGA CGTGAAAGAG CCAGGGGGGG AAGTCGTGAA
      CTCCGGCGGC CCCATCTCCT GCACTTTCTC GGTCCCCCCC TTCAGCACTT
7551  AGAGCCAGGG GGAGAGGTCG TGGACGTGGA GAAAAGAGGC CCAGGAGTCC
      TCTCGGTCCC CCTCTCCAGC ACCTGCACCT CTTTTCTCCG GGTCCTCAGG
7601  CAGTAGTCAG TCATCATCAT CCGGGTCTCC ACCGCGCAGG CCCCCTCCAG
      GTCATCAGTC AGTAGTAGTA GGCCCAGAGG TGGCGCGTCC GGGGGAGGTC
7651  GTAGAAGGCC ATTTTTCCAC CCTGTAGGGG AAGCCGATTA TTTTGAATAC
      CATCTTCCGG TAAAAAGGTG GGACATCCCC TTCGGCTAAT AAAACTTATG
7701  CACCAAGAAG GTGGCCCAGA TGGTGAGCCT GACGTGCCCC CGGGAGCGAT
      GTGGTTCTTC CACCGGGTCT ACCACTCGGA CTGCACGGGG GCCCTCGCTA
7751  AGAGCAGGGC CCCGCAGATG ACCCAGGAGA AGGCCCAAGC ACTGGACCCC
      TCTCGTCCCG GGGCGTCTAC TGGGTCCTCT TCCGGGTTCG TGACCTGGGG
7801  GGGGTCAGGG TGATGGAGGC AGGCGCAAAA AAGGAGGGTG GTTTGGAAAG
      CCCCAGTCCC ACTACCTCCG TCCGCGTTTT TTCCTCCCAC CAAACCTTTC
7851  CATCGTGGTC AAGGAGGTTC CAACCCGAAA TTTGAGAACA TTGCAGAAGG
      GTAGCACCAG TTCCTCCAAG GTTGGGCTTT AAACTCTTGT AACGTCTTCC
7901  TTTAAGAGCT CTCCTGGCTA GGAGTCACGT AGAAAGGACT ACCGACGAAG
      AAATTCTCGA GAGGACCGAT CCTCAGTGCA TCTTTCCTGA TGGCTGCTTC
7951  GAACTTGGGT CGCCGGTGTG TTCGTATATG GAGGTAGTAA GACCTCCCTT
      CTTGAACCCA GCGGCCACAC AAGCATATAC CTCCATCATT CTGGAGGGAA
8001  TACAACCTAA GGCGAGGAAC TGCCCTTGCT ATTCCACAAT GTCGTCTTAC
      ATGTTGGATT CCGCTCCTTG ACGGGAACGA TAAGGTGTTA CAGCAGAATG
8051  ACCATTGAGT CGTCTCCCCT TTGGAATGGC CCCTGGACCC GGCCCACAAC
      TGGTAACTCA GCAGAGGGGA AACCTTACCG GGGACCTGGG CCGGGTGTTG
8101  CTGGCCCGCT AAGGGAGTCC ATTGTCTGTT ATTTCATGGT CTTTTTACAA
      GACCGGGCGA TTCCCTCAGG TAACAGACAA TAAAGTACCA GAAAAATGTT
8151  ACTCATATAT TTGCTGAGGT TTTGAAGGAT GCGATTAAGG ACCTTGTTAT
      TGAGTATATA AACGACTCCA AAACTTCCTA CGCTAATTCC TGGAACAATA
8201  GACAAAGCCC GCTCCTACCT GCAATATCAG GGTGACTGTG TGCAGCTTTG
      CTGTTTCGGG CGAGGATGGA CGTTATAGTC CCACTGACAC ACGTCGAAAC
8251  ACGATGGAGT AGATTTGCCT CCCTGGTTTC CACCTATGGT GGAAGGGGCT
      TGCTACCTCA TCTAAACGGA GGGACCAAAG GTGGATACCA CCTTCCCCGA
8301  GCCGCGGAGG GTGATGACGG AGATGACGGA GATGAAGGAG GTGATGGAGA
      CGGCGCCTCC CACTACTGCC TCTACTGCCT CTACTTCCTC CACTACCTCT
8351  TGAGGGTGAG GAAGGGCAGG AGTGATGTAA CTTGTTAGGA GACGCCCTCA
      ACTCCCACTC CTTCCCGTCC TCACTACATT GAACAATCCT CTGCGGGAGT
8401  ATCGTATTAA AAGCCGTGTA TTCCCCCGCA CTAAAGAATA AATCCCAGT
      TAGCATAATT TTCGGCACAT AAGGGGGCGT GATTTCTTAT TTAGGGGTCA
8451  AGACATCATG CGTGCTGTTG GTGTATTTCT GGCCATCTGT CTTGTCACCA
      TCTGTAGTAC GCACGACAAC CACATAAAGA CCGGTAGACA GAACAGTGGT
8501  TTTTCGTCCT CCCAACATGG GGCAATTGGG CATACCCATG TTGTCACGTC
      AAAAGCAGGA GGGTTGTACC CCGTTAACCC GTATGGGTAC AACAGTGCAG
8551  ACTCAGCTCC GCGCTCAACA CCTTCTCGCG TTGGAAAACA TTAGCGACAT
      TGAGTCGAGG CGCGAGTTGT GGAAGAGCGC AACCTTTTGT AATCGCTGTA
8601  TTACCTGGTG AGCAATCAGA CATGCGACGG CTTTAGCCTG GCCTCCTTAA
      AATGGACCAC TCGTTAGTCT GTACGCTGCC GAAATCGGAC CGGAGGAATT
```

FIGURE 11-8

8

```
8651   ATTCACCTAA GAATGGGAGC AACCAGCATG CAGGAAAAGG ACAAGCAGCG
       TAAGTGGATT CTTACCCTCG TTGGTCGTAC GTCCTTTTCC TGTTCGTCGC
8701   AAAATTCACG CCCCCTTGGG AGGTGGCGGC ATATGCAAAG GATAGCACTC
       TTTTAAGTGC GGGGGAACCC TCCACCGCCG TATACGTTTC CTATCGTGAG
8751   CCACTCTACT ACTGGGTATC ATATGCTGAC TGTATATGCA TGAGGATAGC
       GGTGAGATGA TGACCCATAG TATACGACTG ACATATACGT ACTCCTATCG
8801   ATATGCTACC CGGATACAGA TTAGGATAGC ATATACTACC CAGATATAGA
       TATACGATGG GCCTATGTCT AATCCTATCG TATATGATGG GTCTATATCT
8851   TTAGGATAGC ATATGCTACC CAGATATAGA TTAGGATAGC CTATGCTACC
       AATCCTATCG TATACGATGG GTCTATATCT AATCCTATCG GATACGATGG
8901   CAGATATAAA TTAGGATAGC ATATACTACC CAGATATAGA TTAGGATAGC
       GTCTATATTT AATCCTATCG TATATGATGG GTCTATATCT AATCCTATCG
8951   ATATGCTACC CAGATATAGA TTAGGATAGC CTATGCTACC CAGATATAGA
       TATACGATGG GTCTATATCT AATCCTATCG GATACGATGG GTCTATATCT
9001   TTAGGATAGC ATATGCTACC CAGATATAGA TTAGGATAGC ATATGCTATC
       AATCCTATCG TATACGATGG GTCTATATCT AATCCTATCG TATACGATAG
9051   CAGATATTTG GGTAGTATAT GCTACCCAGA TATAAATTAG GATAGCATAT
       GTCTATAAAC CCATCATATA CGATGGGTCT ATATTTAATC CTATCGTATA
9101   ACTACCCTAA TCTCTATTAG GATAGCATAT GCTACCCGGA TACAGATTAG
       TGATGGGATT AGAGATAATC CTATCGTATA CGATGGGCCT ATGTCTAATC
9151   GATAGCATAT ACTACCCAGA TATAGATTAG GATAGCATAT GCTACCCAGA
       CTATCGTATA TGATGGGTCT ATATCTAATC CTATCGTATA CGATGGGTCT
9201   TATAGATTAG GATAGCCTAT GCTACCCAGA TATAAATTAG GATAGCATAT
       ATATCTAATC CTATCGGATA CGATGGGTCT ATATTTAATC CTATCGTATA
9251   ACTACCCAGA TATAGATTAG GATAGCATAT GCTACCCAGA TATAGATTAG
       TGATGGGTCT ATATCTAATC CTATCGTATA CGATGGGTCT ATATCTAATC
9301   GATAGCCTAT GCTACCCAGA TATAGATTAG GATAGCATAT GCTATCCAGA
       CTATCGGATA CGATGGGTCT ATATCTAATC CTATCGTATA CGATAGGTCT
9351   TATTTGGGTA GTATATGCTA CCCATGGCAA CATTAGCCCA CCGTGCTCTC
       ATAAACCCAT CATATACGAT GGGTACCGTT GTAATCGGGT GGCACGAGAG
9401   AGCGACCTCG TGAATATGAG GACCAACAAC CCTGTGCTTG GCGCTCAGGC
       TCGCTGGAGC ACTTATACTC CTGGTTGTTG GGACACGAAC CGCGAGTCCG
9451   GCAAGTGTGT GTAATTTGTC CTCCAGATCG CAGCAATCGC GCCCCTATCT
       CGTTCACACA CATTAAACAG GAGGTCTAGC GTCGTTAGCG CGGGGATAGA
9501   TGGCCCGCCC ACCTACTTAT GCAGGTATTC CCCGGGGTGC CATTAGTGGT
       ACCGGGCGGG TGGATGAATA CGTCCATAAG GGGCCCCACG GTAATCACCA
9551   TTTGTGGGCA AGTGGTTTGA CCGCAGTGGT TAGCGGGGTT ACAATCAGCC
       AAACACCCGT TCACCAAACT GGCGTCACCA ATCGCCCCAA TGTTAGTCGG
9601   AAGTTATTAC ACCCTTATTT TACAGTCCAA AACCGCAGGG CGGCGTGTGG
       TTCAATAATG TGGGAATAAA ATGTCAGGTT TTGGCGTCCC GCCGCACACC
9651   GGGCTGACGC GTGCCCCCAC TCCACAATTT CAAAAAAAAG AGTGGCCACT
       CCCGACTGCG CACGGGGGTG AGGTGTTAAA GTTTTTTTTC TCACCGGTGA
9701   TGTCTTTGTT TATGGGCCCC ATTGGCGTGG AGCCCCGTTT AATTTTCGGG
       ACAGAAACAA ATACCCGGGG TAACCGCACC TCGGGGCAAA TTAAAAGCCC
9751   GGTGTTAGAG ACAACCAGTG GAGTCCGCTG CTGTCGGCGT CCACTCTCTT
       CCACAATCTC TGTTGGTCAC CTCAGGCGAC GACAGCCGCA GGTGAGAGAA
```

FIGURE 11-9

9

```
 9801   TCCCCTTGTT   ACAAATAGAG   TGTAACAACA   TGGTTCACCT   GTCTTGGTCC
        AGGGGAACAA   TGTTTATCTC   ACATTGTTGT   ACCAAGTGGA   CAGAACCAGG
 9851   CTGCCTGGGA   CACATCTTAA   TAACCCCAGT   ATCATATTGC   ACTAGGATTA
        GACGGACCCT   GTGTAGAATT   ATTGGGGTCA   TAGTATAACG   TGATCCTAAT
 9901   TGTGTTGCCC   ATAGCCATAA   ATTCGTGTGA   GATGGACATC   CAGTCTTTAC
        ACACAACGGG   TATCGGTATT   TAAGCACACT   CTACCTGTAG   GTCAGAAATG
 9951   GGCTTGTCCC   CACCCCATGG   ATTTCTATTG   TTAAAGATAT   TCAGAATGTT
        CCGAACAGGG   GTGGGGTACC   TAAAGATAAC   AATTTCTATA   AGTCTTACAA
10001   TCATTCCTAC   ACTAGTATTT   ATTGCCCAAG   GGGTTTGTGA   GGGTTATATT
        AGTAAGGATG   TGATCATAAA   TAACGGGTTC   CCCAAACACT   CCCAATATAA
10051   GGTGTCATAG   CACAATGCCA   CCACTGAACC   CCCCGTCCAA   ATTTTATTCT
        CCACAGTATC   GTGTTACGGT   GGTGACTTGG   GGGGCAGGTT   TAAAATAAGA
10101   GGGGGCGTCA   CCTGAAACCT   TGTTTTCGAG   CACCTCACAT   ACACCTTACT
        CCCCCGCAGT   GGACTTTGGA   ACAAAAGCTC   GTGGAGTGTA   TGTGGAATGA
10151   GTTCACAACT   CAGCAGTTAT   TCTATTAGCT   AAACGAAGGA   GAATGAAGAA
        CAAGTGTTGA   GTCGTCAATA   AGATAATCGA   TTTGCTTCCT   CTTACTTCTT
10201   GCAGGCGAAG   ATTCAGGAGA   GTTCACTGCC   CGCTCCTTGA   TCTTCAGCCA
        CGTCCGCTTC   TAAGTCCTCT   CAAGTGACGG   GCGAGGAACT   AGAAGTCGGT
10251   CTGCCCTTGT   GACTAAAATG   GTTCACTACC   CTCGTGGAAT   CCTGACCCCA
        GACGGGAACA   CTGATTTTAC   CAAGTGATGG   GAGCACCTTA   GGACTGGGGT
10301   TGTAAATAAA   ACCGTGACAG   CTCATGGGGT   GGGAGATATC   GCTGTTCCTT
        ACATTTATTT   TGGCACTGTC   GAGTACCCCA   CCCTCTATAG   CGACAAGGAA
10351   AGGACCCTTT   TACTAACCCT   AATTCGATAG   CATATGCTTC   CCGTTGGGTA
        TCCTGGGAAA   ATGATTGGGA   TTAAGCTATC   GTATACGAAG   GGCAACCCAT
10401   ACATATGCTA   TTGAATTAGG   GTTAGTCTGG   ATAGTATATA   CTACTACCCG
        TGTATACGAT   AACTTAATCC   CAATCAGACC   TATCATATAT   GATGATGGGC
10451   GGAAGCATAT   GCTACCCGTT   TAGGGTTAAC   AAGGGGGCCT   TATAAACACT
        CCTTCGTATA   CGATGGGCAA   ATCCCAATTG   TTCCCCCGGA   ATATTTGTGA
10501   ATTGCTAATG   CCCTCTTGAG   GGTCCGCTTA   TCGGTAGCTA   CACAGGCCCC
        TAACGATTAC   GGGAGAACTC   CCAGGCGAAT   AGCCATCGAT   GTGTCCGGGG
10551   TCTGATTGAC   GTTGGTGTAG   CCTCCCGTAG   TCTTCCTGGG   CCCCTGGGAG
        AGACTAACTG   CAACCACATC   GGAGGGCATC   AGAAGGACCC   GGGGACCCTC
10601   GTACATGTCC   CCCAGCATTG   GTGTAAGAGC   TTCAGCCAAG   AGTTACACAT
        CATGTACAGG   GGGTCGTAAC   CACATTCTCG   AAGTCGGTTC   TCAATGTGTA
10651   AAAGGCAATG   TTGTGTTGCA   GTCCACAGAC   TGCAAAGTCT   GCTCCAGGAT
        TTTCCGTTAC   AACACAACGT   CAGGTGTCTG   ACGTTTCAGA   CGAGGTCCTA
10701   GAAAGCCACT   CAGTGTTGGC   AAATGTGCAC   ATCCATTTAT   AAGGATGTCA
        CTTTCGGTGA   GTCACAACCG   TTTACACGTG   TAGGTAAATA   TTCCTACAGT
10751   ACTACAGTCA   GAGAACCCCT   TTGTGTTTGG   TCCCCCCCCG   TGTCACATGT
        TGATGTCAGT   CTCTTGGGGA   AACACAAACC   AGGGGGGGGC   ACAGTGTACA
10801   GGAACAGGGC   CCAGTTGGCA   AGTTGTACCA   ACCAACTGAA   GGGATTACAT
        CCTTGTCCCG   GGTCAACCGT   TCAACATGGT   TGGTTGACTT   CCCTAATGTA
10851   GCACTGCCCC   GCGAAGAAGG   GGCAGAGATG   CCGTAGTCAG   GTTTAGTTCG
        CGTGACGGGG   CGCTTCTTCC   CCGTCTCTAC   GGCATCAGTC   CAAATCAAGC
10901   TCCGGCGGCG   GGGCTCTAGA   GTCGACCGGT   CATGGCTGCG   CCCCGACACC
        AGGCCGCCGC   CCCGAGATCT   CAGCTGGCCA   GTACCGACGC   GGGGCTGTGG
```
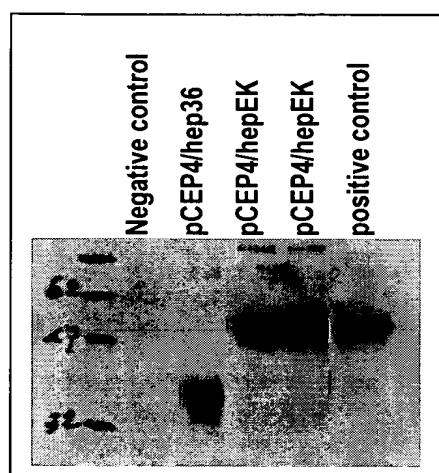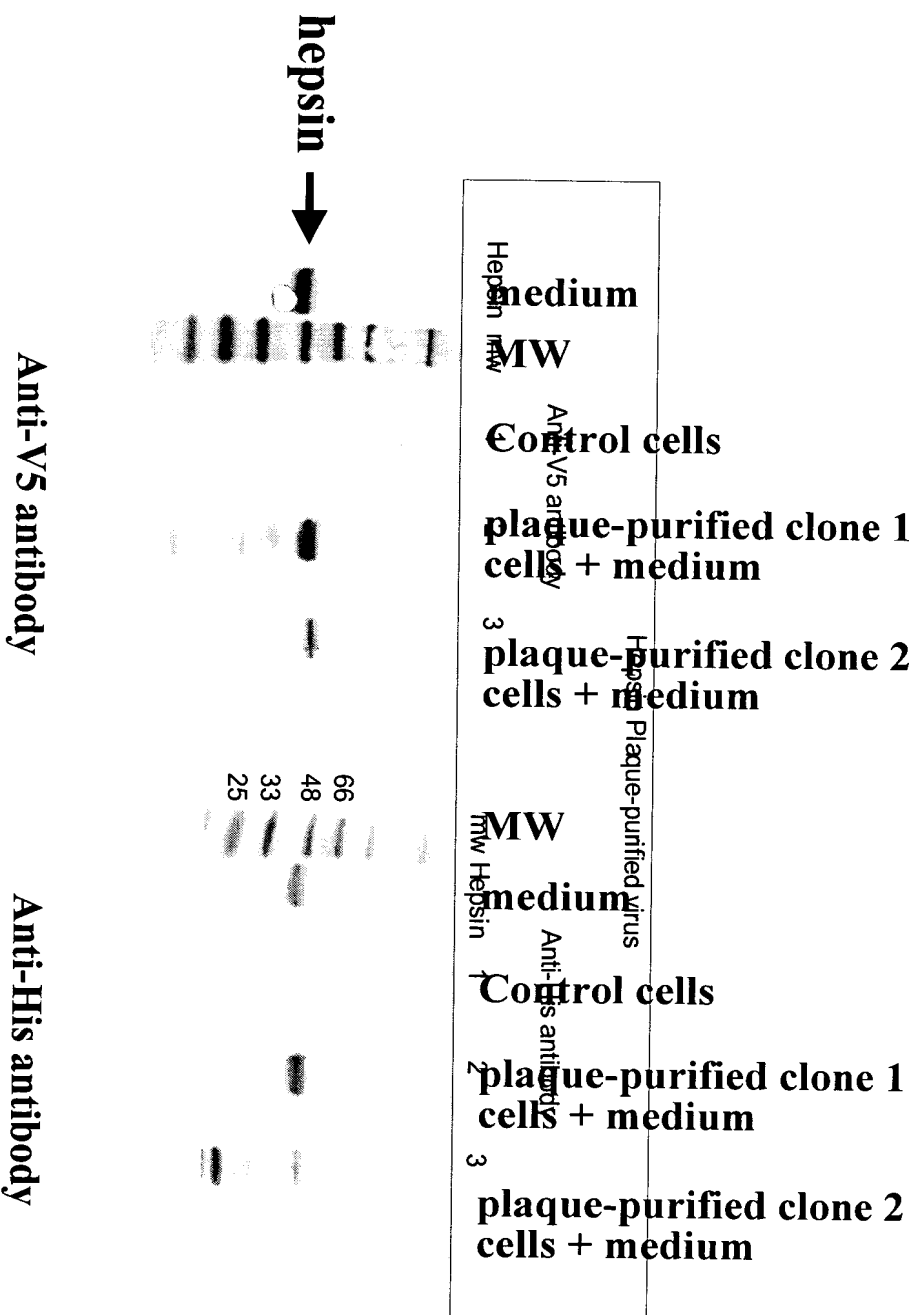
FIGURE 11-10

10951 CGCCAACACC CGCTGACGCG CCCTGACGGG CTTGTCTGCT CCCGGCATCC
      GCGGTTGTGG CGACTGCGC GGGACTGCCC GAACAGACGA GGGCCGTAGG

11001 GCTTACAGAC AAGCTGTGAC CGTCTCCGGG AGCTGCATGT GTCAGAGGTT
      CGAATGTCTG TTCGACACTG GCAGAGGCCC TCGACGTACA CAGTCTCCAA

11051 TTCACCGTCA TCACCGAAAC GCGCGAGGCA GCCGGATCAT AATCAGCCAT
      AAGTGGCAGT AGTGGCTTTG CGCGCTCCGT CGGCCTAGTA TTAGTCGGTA

11101 ACCACATTTG TAGAGGTTTT ACTTGCTTTA AAAAACCTCC CCACCTCCCC
      TGGTGTAAAC ATCTCCAAAA TGAACGAAAT TTTTTGGAGG GGTGGAGGGG

11151 CTGAACCTGA AACATAAAAT GAATGCAATT GTTGTTGTTA ACTTGTTTAT
      GACTTGGACT TTGTATTTTA CTTACGTTAA CAACAACAAT TGAACAAATA

11201 TGCAGCTTAT AATGGTTACA AATAAAGCAA TAGCATCACA AATTTCACAA
      ACGTCGAATA TTACCAATGT TTATTTCGTT ATCGTAGTGT TTAAAGTGTT

11251 ATAAAGCATT TTTTTCACTG CATTCTAGTT GTGGTTTGTC CAAACTCATC
      TATTTCGTAA AAAAAGTGAC GTAAGATCAA CACCAAACAG GTTTGAGTAG

11301 AATGTATCTT ATCATGTCTG GATCCCACGT GCAGGCGGGG AGGCGGCCCA
      TTACATAGAA TAGTACAGAC CTAGGGTGCA CGTCCGCCCC TCCGCCGGGT

11351 AAGGGAGATC CGACTCGTCT GAGGGCGAAG GCGAAGACGC GGAAGAGGCC
      TTCCCTCTAG GCTGAGCAGA CTCCCGCTTC CGCTTCTGCG CCTTCTCCGG

11401 GCAGAGCCGG CAGCAGGCCG CGGGAAGGAA GGTCCGCTGG ATTGAGGGCC
      CGTCTCGGCC GTCGTCCGGC GCCCTTCCTT CCAGGCGACC TAACTCCCGG

11451 GAAGGGACGT AGCAGAAGGA CGTCCCGCGC AGAATCCAGG TGGCAACACA
      CTTCCCTGCA TCGTCTTCCT GCAGGGCGCG TCTTAGGTCC ACCGTTGTGT

11501 GGCGAGCAGC CAAGGAAAGG ACGATGATTT CCCCGACAAC ACCACGGAAT
      CCGCTCGTCG GTTCCTTTCC TGCTACTAAA GGGGCTGTTG TGGTGCCTTA

11551 TGTCAGTGCC CAACAGCCGA GCCCCTGTCC AGCAGCGGGC AAGGCAGGCG
      ACAGTCACGG GTTGTCGGCT CGGGGACAGG TCGTCGCCCG TTCCGTCCGC

11601 GCGATGAGTT CCGCCGTGGC AATAGGGAGG GGGAAAGCGA AAGTCCCGGA
      CGCTACTCAA GGCGGCACCG TTATCCCTCC CCCTTTCGCT TTCAGGGCCT

11651 AAGGAGCTGA CAGGTGGTGG CAATGCCCCA ACCAGTGGGG GTTGCGTCAG
      TTCCTCGACT GTCCACCACC GTTACGGGGT TGGTCACCCC CAACGCAGTC

11701 CAAACACAGT GCACCACG CCACGTTGCC TGACAACGGG CCACAACTCC
      GTTTGTGTCA CGTGTGGTGC GGTGCAACGG ACTGTTGCCC GGTGTTGAGG

11751 TCATAAAGAG ACAGCAACCA GGATTTATAC AAGGAGGAGA AAATGAAAGC
      AGTATTTCTC TGTCGTTGGT CCTAAATATG TTCCTCCTCT TTTACTTTCG

11801 CATACGGGAA GCAATAGCAT GATACAAAGG CATTAAAGCA GCGTATCCAC
      GTATGCCCTT CGTTATCGTA CTATGTTTCC GTAATTTCGT CGCATAGGTG

11851 ATAGCGTAAA AGGAGCAACA TAGTTAAGAA TACCAGTCAA TCTTTCACAA
      TATCGCATTT TCCTCGTTGT ATCAATTCTT ATGGTCAGTT AGAAAGTGTT

11901 ATTTTGTAAT CCAGAGGTTG ATTC
      TAAAACATTA GGTCTCCAAC TAAG

FIGURE 11-11

Figure 12: Western Blot of Hepsin expressed in 293EBNA cells. 4 days after transfection, conditioned medium was harvested. 10 µl of conditioned medium was loaded on a 10% Tris-Glycine gel. Untreated 293EBNA cells served as negative control, 5 ng of purified hepsin ED/EK (expressed from insect cells) served as positive control. After electrophoresis, proteins were transferred onto Nitrocellulose membrane at 100 V for 1h. The membrane was then probed with anti-V5 antibody at 37C° for 1h after blocking with 5% dry milk in TTBS buffer for 1h.
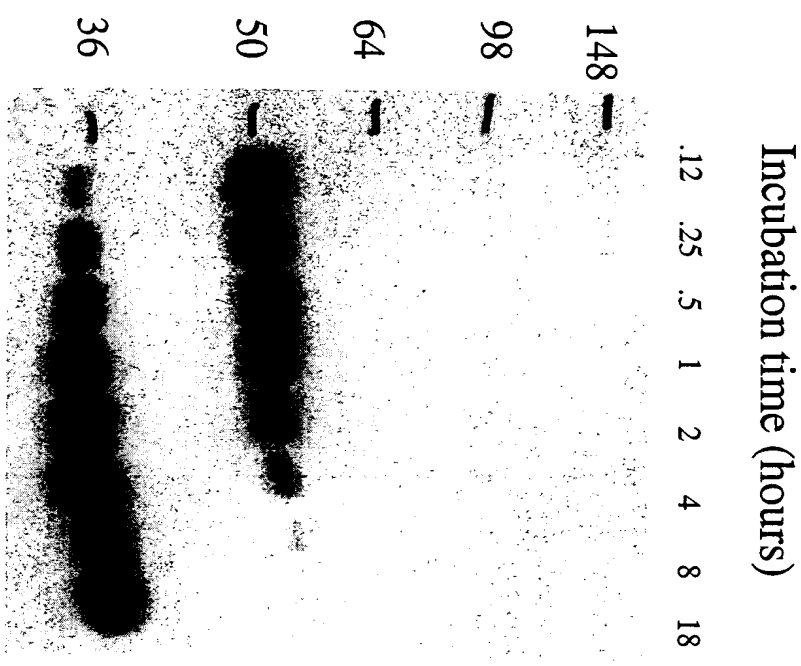
hepsin

Anti-V5 antibody   Anti-His antibody

66
48
33
25

medium
MW
Control cells
plaque-purified clone 1 cells + medium
plaque-purified clone 2 cells + medium
MW
medium
Control cells
plaque-purified clone 1 cells + medium
plaque-purified clone 2 cells + medium

Fig. 13. Western analysis of recombinant soluble hepsin. Recombinant soluble hepsin was expressed in baculovirus infected insect cells. Western analysis using an anti-V5 antibody (left panel) and an anti-His antibody (right panel) showed that recombinant soluble hepsin was present in cultured medium and cell pellets.
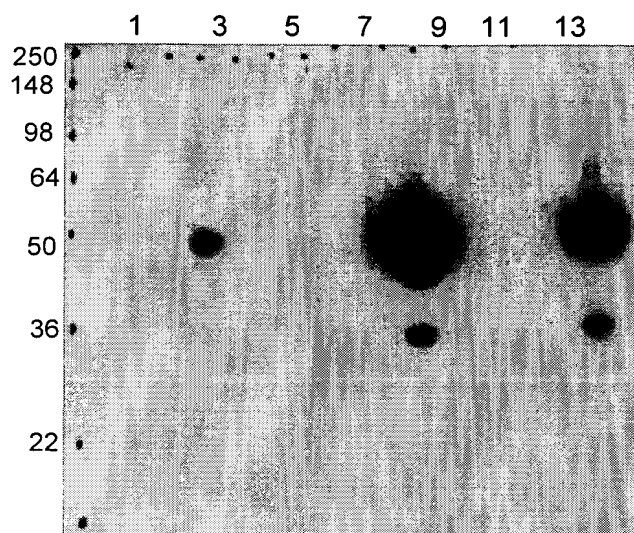
Figure 14: Enterokinase (EKMax)processing of Hepsin -ED-EK protein to generate active enzyme

3.1 units EKMax / ml

Figure 15A: Western Slot Blot on the hybridomas 11C1 & 47A5 are WB positive against Hepsin-ED-EK protein.

Blot 1		
slot #	I.D.	CODE
1	1	2F10
2	2	9H6
3	3	11C1
4	4	16A2
5	5	27A1
6	6	30C10
7	7	30H12
8	8	35D4
9	9	47A5
10	10	80F6
11	Blank	
12	Blank	
13	Blank	
14	Mouse Poly clonal	

Figure 15B: Western Slot Blot on the hybridomas 31C1 and 38E2 are WB positive against Hepsin-ED-EK protein.

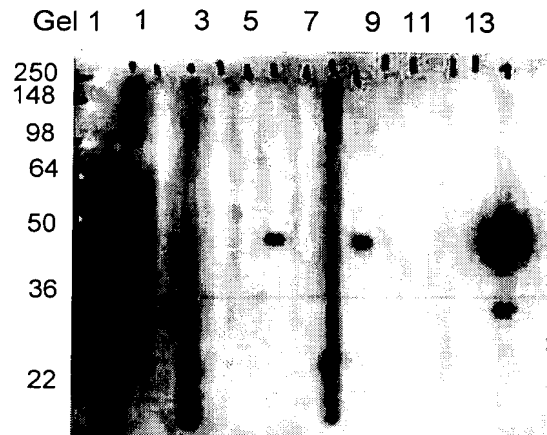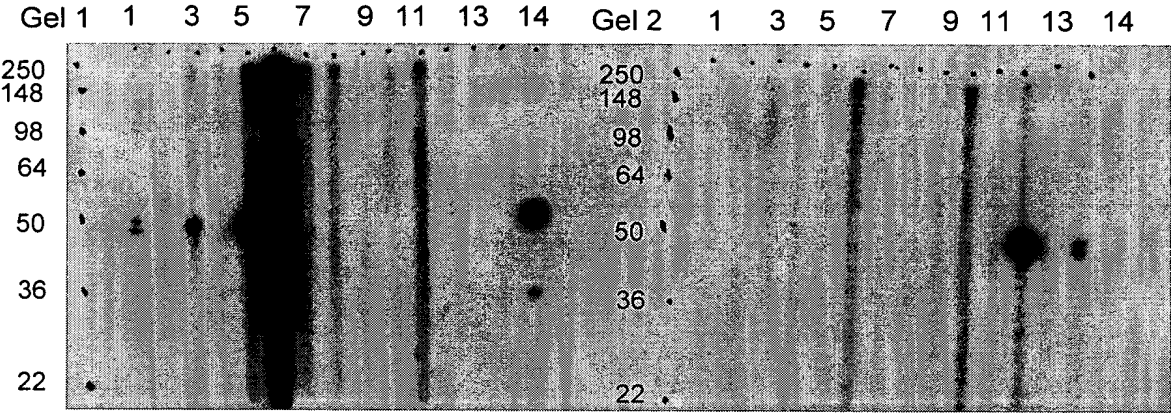| Blot 1		
slot #	I.D.	CODE
1	1	7H3
2	2	12A2
3	3	16A3
4	4	20D1
5	5	22A3
6	6	31C1
7	7	31H2
8	8	36B2
9	9	38 E2
10	10	39D6
11	11	41F7
12	12	42 E3
13	Blank	
14	Mouse Poly clonal	

Figure 15C: Western Slot Blot on the hybridomas 37G10, 46D12 & 14C7 are WB positive against Hepsin-ED-EK protein.

Gel 1

Slot	Sample	Dilution
1	Rabbit Polyclonal –Cayman Chem	1:500
2	Media	neat
3	37G10	"
4	94A7	"
5	46D12	"
6	103E3	"
7	40F1	"
8	103H12	"
9	93D1	"
10	10C2	"
11	102F2	"
12	83E11	"
13	Blank	
14	Blank	
15	Protein - ♂ mouse	1:5000

Gel 2

Slot	Sample	Dilution
1	Protein - ♂ mouse	1:5000
2	Media	neat
3	27E7	"
4	92A7	"
5	91A4	"
6	99B11	"
7	94C7	"
8	91A1	"
9	14H11	"
10	74C7	"
11	72H6	"
12	14C7	"
13	Blank	
14	Rabbit Polyclonal –Cayman Chem	1:500

Figure 15D: Western Slot Blot on the hybridomas 14C7 & 72H6 are WB positive against Hepsin-ED-EK protein.
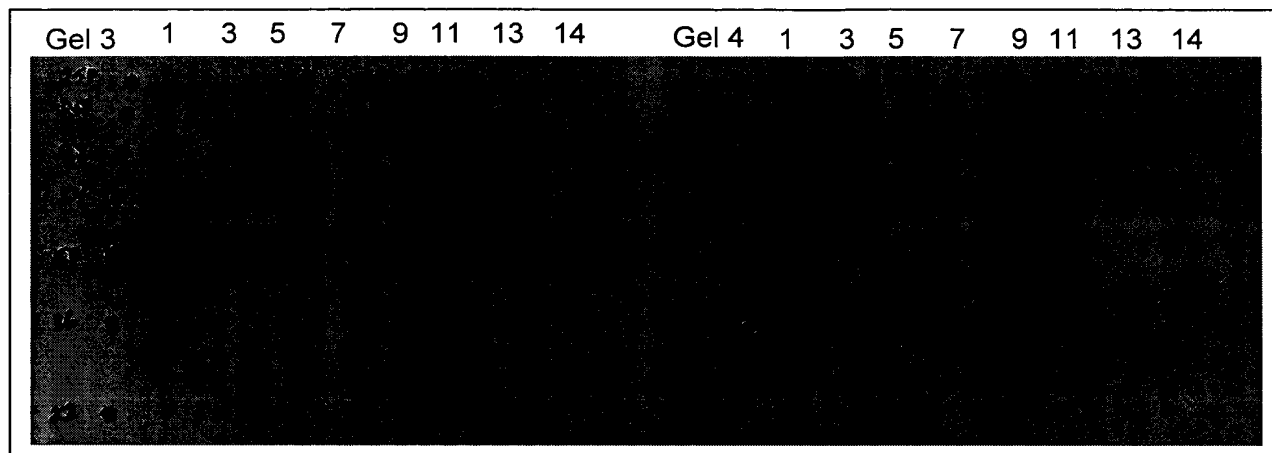
Gel 3

Slot	Sample	Dilution
1	Protein - ♂ mouse	1:5000
2	Media	neat
3	84G6	"
4	84H2	"
5	51F8	"
6	15b11	"
7	90E6	"
8	72H6	"
9	85A4	"
10	87C2	"
11	3G11	"
12	53E11	"
13	91H4	"
14	53C7	"
15	88C7	"

Gel 4

Slot	Sample	Dilution
1	Protein - ♂ mouse	1:5000
2	Media	neat
3	75H3	"
4	98B4	"
5	91C9	"
6	53D9	"
7	80G6	"
8	95F3	"
9	14C7	"
10	80H10	"
11	92A9	"
12	81C8	"
13	96B6	"
14	18I2	"

FIG 16A: Human Prostate Tumor Tumor Tissue
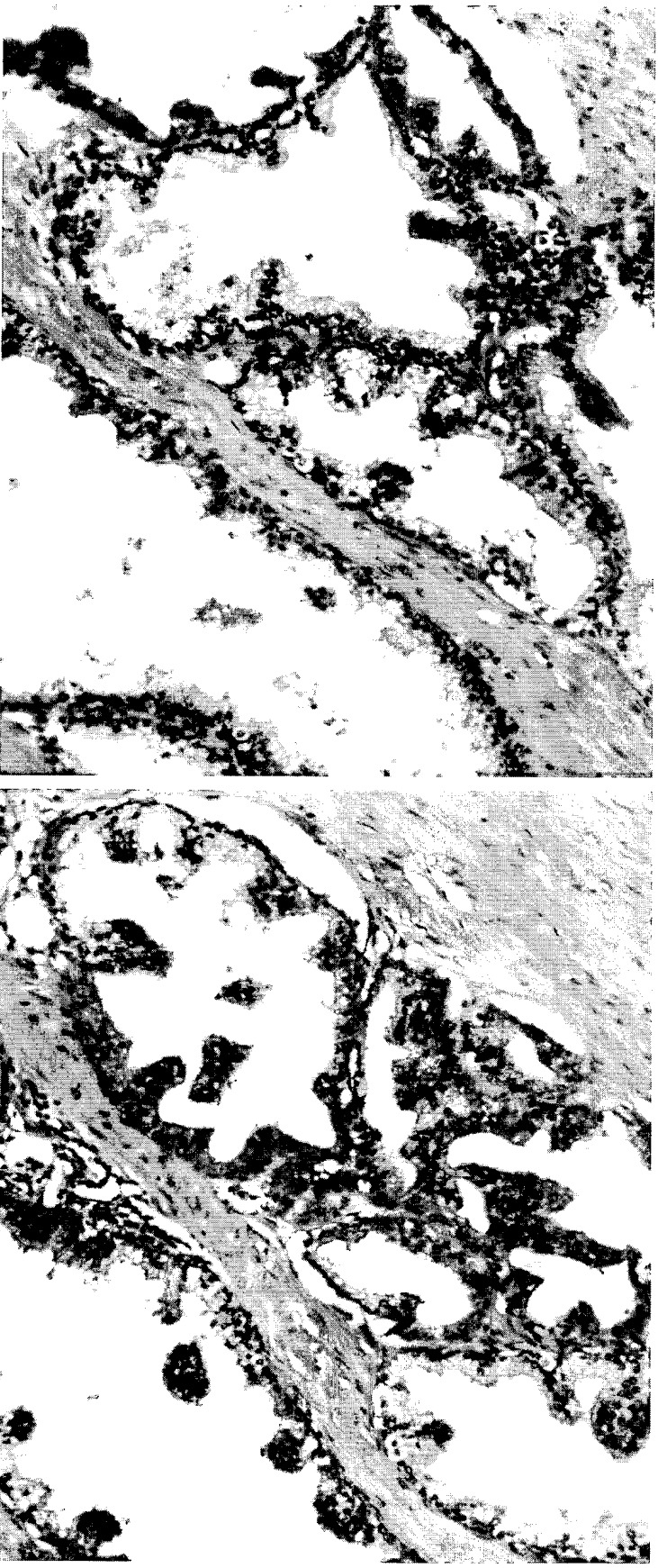
Control
Anti-Hepsin mouse
polyclonal preimmun
1/500 (9)

Anti-Hepsin mouse
polyclonal immun
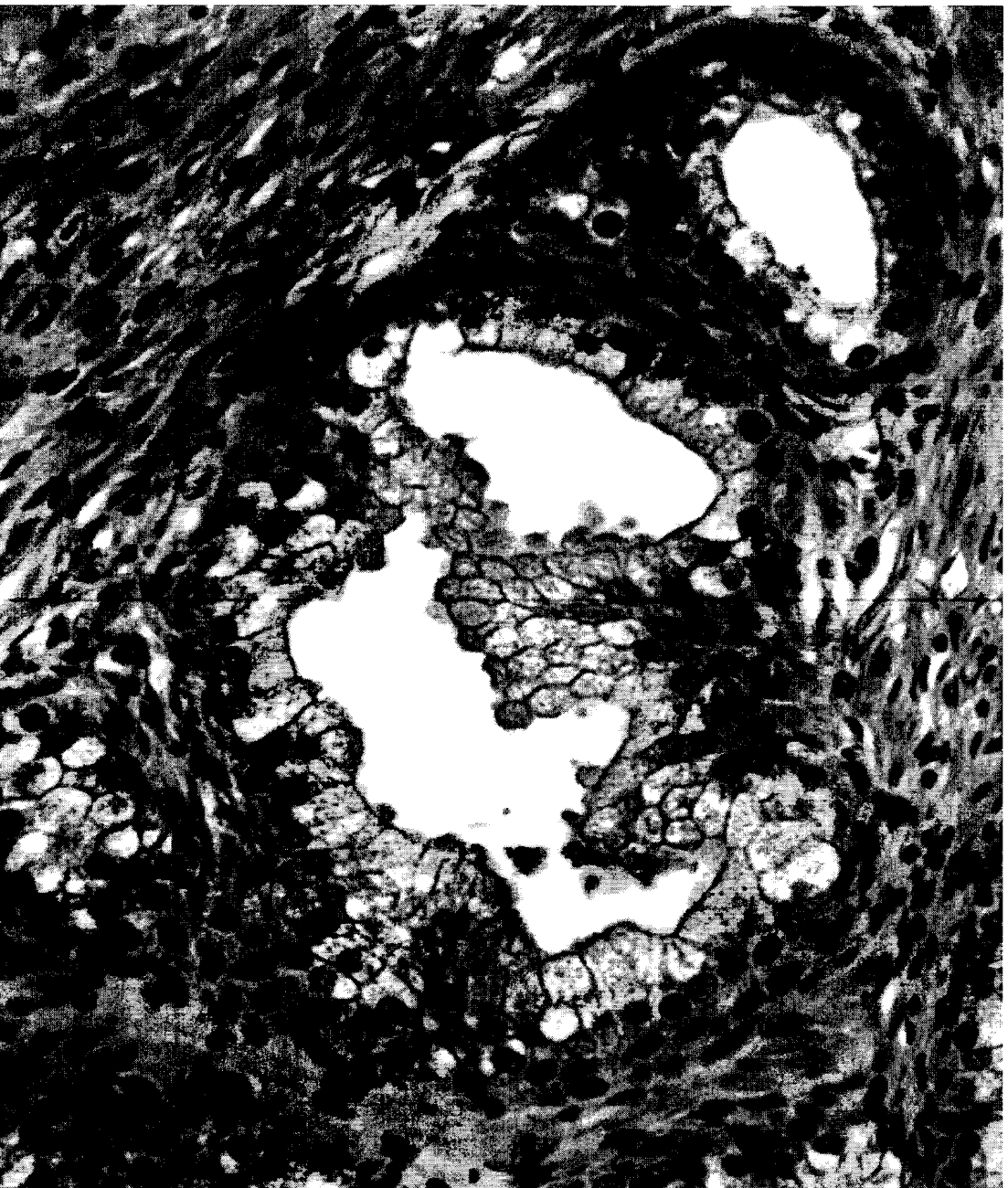1/500 (10)

FIG 16B: **Human Prostate Tumor Tissue**

Control medium

Anti-Hepsin monoclonal ab (medium)

FIG 16C: Human Prostate Tumor Tissue

Anti-Hepsin monoclonal ab

Figure 17

Human wild-type hepsin amino acid sequence:

```
maqkeggrtv pccsrpkvaa ltagtllllt aigaaswaiv avllrsdqep lypvqvssad 61
arlmvfdkte gtwrllcssr snarvaglsc eemgflralt hseldvrtag angtsgffcv 121
degrlphtqr llevisvcdc prgrflaaic qdcgrrklpv drivggrdts lgrwpwqvsl 181
rydgahlcgg sllsgdwvlt aahcfpernr vlsrwrvfag avaqasphgl qlgvqavvyh 241
ggylpfrdpn seensndial vhlssplplt eyiqpvclpa agqalvdgki ctvtgwgntq 301
yygqqagvlq earvpiisnd vcngadfygn qikpkmfcag ypeggidacq gdsggpfvce 361
dsisrtprwr lcgivswgtg calaqkpgvy tkvsdfrewi fqaikthsea sgmvtql
```

the cytoplasmic domain: Met1 to Lys17
the transmembrane domain: Val18 to Leu44
the ectodomain: Arg45 to Leu417

Figure 18

Hep-ED-EK structure (modified soluble hepsin with substitute activation sequence)

```
  1 RSDQEPLYPV QVSSADARLM VFDKTEGTWR LLCSSRSNAR VAGLSCEEMG FLRALTHSEL
 61 DVRTAGANGT SGFFCVDEGR LPHTQRLLEV ISVCDCPRGR FLAAICQDCG RRKLPVDDDD
121 KIVGGRDTSL GRWPWQVSLR YDGAHLCGGS LLSGDWVLTA AHCFPERNRV LSRWRVFAGA
181 VAQASPHGLQ LGVQAVVYHG GYLPFRDPNS EENSNDIALV HLSSPLPLTE YIQPVCLPAA
241 GQALVDGKIC TVTGWGNTQY YGQQAGVLQE ARVPIISNDV CNGADFYGNQ IKPKMFCAGY
301 PEGGIDACQG DSGGPFVCED SISRTPRWRL CGIVSWGTGC ALAQKPGVYT KVSDFREWIF
361 QAIKTHSEAS GMVTQLEFGK PIPNPLLGLD STRTGHHHHH H*
```

Cytoplasmic domain: absent
Transmembrane domain: absent
Modified activation domain:117-121 (underlined)
V5 and 6-His Tag: 377-401

Figure 19

Antibody Neutralization

Relative Rate %

Legend:
■ Hepsin tf BHK
▦ Hepsin tf BHK no F7
▨ Hepsin tf BHK
▢ BHK parental

X-axis categories: IgG, no Ab (+) F7, no Ab (-) F7, 1A12, 90C1, 85B11, 40F1, 7H3, 94A7, 27E3